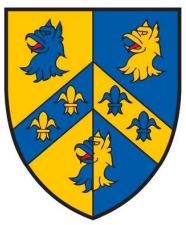
Regulation of Molecular Subtypes of Pancreatic Cancer by TET2 and 5'hmC



Michael Eyres

Trinity College

University of Oxford



A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2019

Supervisor:
Prof. Eric O'Neill (University of Oxford)



The work presented in this thesis was performed at the Old Road Campus Research Building (ORCRB), Department of Oncology, University of Oxford. This study was funded by Medical Research Council and Oxford Department of Oncology funded Prize Studentship. Additional funding was provided by Precision Panc, Cancer Research UK, the Kadani Foundation and the Sarah and Nadine Pole Scholarship.

To my Family,
for helping me through a decade of University

Abstract

Genomic studies using mRNA datasets have uncovered two molecular subtypes within pancreatic ductal adenocarcinoma (PDAC) tumours with distinct gene programmes and clinical characteristics, raising the possibility of subtype specific therapeutic options. In each of these studies a similar aggressive subtype was identified that was characterized by a loss of endodermal differentiation markers such as GATA6 and PDX1, as well as poorer patient survival, referred to collectively as squamous. Despite the inherent heterogeneity in the molecular basis of PDAC, there is little genetic heterogeneity between different PDAC tumours, even when comparing metastasis to primary tumours. Conversely, there is widespread epigenetic reprogramming association with progression to more aggressive phenotypes, suggesting that aggressive molecular subtypes are likely to be epigenetically driven.

TET-mediated DNA Hydroxymethylation (5'hmC) is dynamically regulated during pancreatic differentiation, but it remains unknown what role 5'hmC during PDAC initiation and progression. In this work the role of 5'hmC during PDAC progression is explored. Loss of 5'hmC correlates with squamous subtypes and decreased expression of GATA6 while re-expression of TET2 increases GATA6 expression. Critically, TET2 can be effectively targeted using Metformin and Vitamin C (Ascorbic Acid). When used in combination, metformin and ascorbic acid can increase TET2 and GATA6 levels *in vitro* and *in vivo*. This suppresses Wnt pathway signalling and causes cells to become less squamous in identity. Thus, the findings presented here suggest that aggressive molecular subtypes can be targeted therapeutically be restoring TET2 activity.

Acknowledgements

Basic research is like shooting an arrow in the air and, where it lands, painting a target; - Homer Burton Adkins

You really need to stop getting distracted and just focus on one project; - Various

Science is intrinsically a collaborative process. The work presented in this thesis and performed throughout my DPhil could not have been achieved without the help and attention of many people, some of who I'd like to take the opportunity to thank here.

First and foremost, I would like to thank my supervisor, Eric O'Neill, for his advice and guidance throughout my DPhil. My work style can be quite hectic and I get easily distracted by new ideas, so having someone who gives me the freedom to explore whilst keeping me on track was crucial for me, and I enjoyed being able to bounce ideas off each another. Furthermore, I would like to thank the rest of the O'Neill lab members: past and present. I received advice from pretty much the whole lab at one point or another, but the help of Dafni, Leanne, Maria, Delia, and Asmita has been particularly invaluable.

My gratitude also goes to the collaborators who were involved in this project. The creation of the Oxford Cohort was only possible through the help of a group of researchers meeting regularly to share their advice and expertise. The initial groundwork for my project was laid by Zenobia D'Costa and Aline Hubert. Adam Burns and Andrew Blake provided crucial QC help and Robert Goldin lent his pathology skills.

Outside of the lab there is a huge network of people who have helped me get through the past four years. To name them all would make for quite a boring list of names, but luckily most of them can be grouped into various Facebook or Whatsapp groups. In Oxford I'd like to thank Greg for all the post-work drinks over the years, as well as Andrew and Dan for kicking me out of bed to go to the gym in the mornings. Additionally, I'd like to thank everyone in the Streitgruppe for being a constant source of entertainment, distractions, and memes.

One of the greatest advantages of spending a decade at University is that it allows you to travel and work all over the world. Without Aron in Melbourne I would never have applied to Oxford in the first place. Furthermore, My friends in the Netherlands, from the original [P]rivate [P]arty to Super SOrrAsxwxywuxw, gave me a chance to recharge and get some distance from my DPhil. Additionally, Chris, Dan, Nick, and Clinton in Cardiff and Patrick, Gilles, Simon, and Justin provided some invaluable weekends away.

I would also like to thank my family. My parents have helped support me through the past 10 years of University life. I promise I'll get a grown up job soon.

Finally, I'd like to thank Tom, who has helped me remain calm throughout my DPhil by being a big enough drama queen for the both of us. Thanks for your encouragement and support, the past four years would have been a lot less fun without you.

Contents

List of Figures x
List of Tables xi

1 Introduction **1**
1.1 Cancer 1
1.1.1 Cancer Stem Cells 3
1.1.2 Pancreatic Ductal Carcinoma (PDAC) 4
1.1.3 Genomic Analyses of Cancer 7
1.1.4 Molecular Subtypes of PDAC 8
1.1.5 Evolution of PDAC Molecular Subtypes 12
1.1.6 GATA6 13
1.1.7 Canonical Wnt signalling and the PDAC Stem Cell Niche . . . 15
1.2 Epigenetics 18
1.2.1 Histones 19
1.2.2 DNA Methylation 20
1.2.3 de-Methylation and Hydroxymethylation (5'hmC) 22
1.2.4 5'hmC Detection Methods 23
1.2.5 5'hmC in the Pancreas 24
1.2.6 TETs 25
1.3 Aims 28

2 Materials and Methods **29**
2.1 Materials 29
2.2 Tables 30
2.3 Generation of Genomic Data from a Cohort of PDAC Patients 33
2.3.1 Ethics and Patient Material 33
2.3.2 mRNA Isolation and Microarrays 34
2.3.3 DNA Isolation 34
2.3.4 Genome Wide Methylation and Hydroxymethylation 34

2.3.5 Ion Torrent Sequencing . . . 36
2.4 Statistical Analysis and Bioinformatics . . . 37
2.4.1 Statistical Tests . . . 37
2.4.2 Clinical Datasets . . . 37
2.4.3 Methylation/Hydroxymethylation Analysis . . . 38
2.4.4 Gene Set Enrichment and Molecular Subtyping . . . 38
2.4.5 iCluster . . . 39
2.5 Molecular Biology and *in vitro* experiments . . . 40
2.5.1 Cell Culture . . . 40
2.5.2 Transfection . . . 40
2.5.3 *in vitro* Drug Treatments . . . 41
2.5.4 Protein Isolation, SDS-Page and Western Blotting . . . 41
2.5.5 DNA Isolation and Dot Blotting . . . 43
2.5.6 RNA Isolation and qPCR . . . 44
2.6 *in vivo* Experiments and Histological Analysis . . . 44
2.6.1 Orthotopic Injection of PDAC Cell Lines into Mice . . . 44
2.6.2 Tissue Processing and Embedding . . . 45
2.6.3 Histological Analysis . . . 45
2.6.4 RNA-seq on Orthotopic Tumours . . . 47

3 Whole Genome 5'hmC Profiling in a Cohort of PDAC Patients 50
3.1 Introduction . . . 50
3.1.1 Bisulfite Conversion of DNA for Detection of Methylation . . . 51
3.1.2 The oxBS method for detection of 5'hmC . . . 51
3.1.3 Aims . . . 52
3.2 Results . . . 54
3.2.1 Generation of Genomic Data from a Cohort of FFPE PDAC Patients - The Oxford Cohort . . . 54
3.2.2 Total 5'hmC levels are reduced during PDAC tumorigenesis . . . 57
3.2.3 Detection of genome wide 5'hmC in FFPE Patient Tissue . . . 59
3.2.4 Pancreatic Tumours are Losing 5'hmC at PDAC Specific Loci 62
3.2.5 Loss of 5'hmC Correlates With Increased 5'mC . . . 62
3.2.6 Integrated Molecular Analysis of PDAC Links 5'hmC with PDAC Molecular Subtype . . . 66
3.3 Discussion . . . 73

4 TET2 Promotes GATA6 and can be Targeted Pharmacologically 76
4.1 Introduction . . . 76
4.1.1 Vitamin C (Ascorbic Acid) Potentates TET Activity . . . 76
4.1.2 Regulation of TET2 by AMPK and Metformin . . . 77
4.1.3 Aims . . . 78
4.2 Results . . . 79
4.2.1 Molecular Classification of Pancreatic Cell Lines . . . 79
4.2.2 Overexpression of TET2 in Aggressive Cells Increases Classical-Pancreatic Subtype Markers . . . 81
4.2.3 Metformin can increase TET2 and GATA6 levels in an AMPK-dependent manner . . . 83
4.2.4 Restoration of 5'hmC by Ascorbic Acid Increases GATA6 . . . 85
4.2.5 Metformin and Ascorbic Acid Act Synergistically to regulate 5'hmC and GATA6 . . . 86
4.3 Discussion . . . 92

5 Pharmacological Restoration of 5'hmC in vivo 95
5.1 Introduction . . . 95
5.1.1 Aims . . . 96
5.2 Results . . . 97
5.2.1 Orthotopic Injection of Human PDAC Cell Lines into Nude Mice 97
5.2.2 Metformin and Ascorbic Acid Increase 5'hmC and GATA6 Levels *in vivo* . . . 101
5.2.3 Restoration of 5'hmC Results in a Partial Switch in Subtype . 104
5.2.4 Metformin and AA Inhibit Wnt Pathway Signalling . . . 106
5.2.5 Metformin/AA Inhibit Tumour Proliferation . . . 111
5.3 Discussion . . . 113

6 SMAD4 Regulates 5'hmC by Promoting TET2 Expression 115
6.1 Introduction . . . 115
6.1.1 SMAD4 Loss in PDAC . . . 115
6.1.2 Aims . . . 117
6.2 Results . . . 118
6.2.1 SMAD4 Promotes TET2 Transcription . . . 118
6.2.2 SMAD4 Regulates Molecular Subtype Markers . . . 122
6.3 Discussion . . . 127

7 Discussion **129**
7.1 Overview of Results . . . 129
7.1.1 5'hmC is a Tumour Suppressor in PDAC . . . 129
7.1.2 TET2 Regulates Molecular Subtype Genes . . . 130
7.1.3 Pharmacological Targeting of TET2 . . . 131
7.2 Potential Mechanisms of TET2 Loss in PDAC . . . 133
7.2.1 Mutations . . . 133
7.2.2 Regulation by other Epigenetic Enzymes . . . 134
7.2.3 Diabetes and Hyperglycemia . . . 135
7.2.4 Hypoxia . . . 136
7.3 Remaining Aspects and Future Work . . . 139
7.3.1 Further Validation of the oxBS Method in FFPE . . . 139
7.3.2 Genome-wide 5'hmC Changes in Treated Cells . . . 139
7.3.3 Determining the Extent of TET2 Involvement in Subtype Changes 140
7.4 Conclusion . . . 140

A Appendix **142**

B Copyright Licences **146**

Bibliography **148**

List of Figures

1.1 The Hallmarks of Cancer 2
1.2 Pancreatic Ductal Carcinoma (PDAC) 6
1.3 Genomic Analyses of PDAC 9
1.4 Molecular Subtypes of PDAC 11
1.5 Loss of GATA6 Promotes Wnt Autonomy and Squamous Subtypes . 16
1.6 Epigenetic Mechanisms 19
1.7 Cytosine Modifications 23
1.8 TET Proteins 26

3.1 Bisulfite Conversion Effects on Cytosine Modifications 52
3.2 The Oxford Cohort 55
3.3 Total 5'hmC Levels Decrease During PDAC Tumorigenesis 58
3.4 Adaption of oxBS Method to FFPE 60
3.5 Detection of Methylation and Hydroxymethylation FFPE 61
3.6 PDAC Tumours Lose 5'hmC at PDAC Specific Loci 63
3.7 Loss of 5'hmC at CpG Shores Correlates with increased 5'mC at CpG Islands 64
3.8 Integrated Genetic and Epigenetic Analysis of PDAC 65
3.9 Differentially Expressed Gene Sets Between Clusters 67
3.10 Molecular Subtyping of the Oxford Cohort 69
3.11 Application of iCluster Methylation Signature to another Cohort . . . 71
3.12 5'hmC Correlates with PDAC Molecular Subtype 72

4.1 Molecular Classification of PDAC Cell Lines 80
4.2 Overexpression of TET2 in Aggressive Cell Lines 82
4.3 Metformin Increases TET2 and GATA6 in PSN1 Cells 84
4.4 Regulation of TET2 and GATA6 by Metformin is AMPK Dependent 85
4.5 Ascorbic Acid Increases 5'hmC and GATA6 Levels 86
4.6 Metformin and Ascorbic Acid Act Synergystically in PSN1 Cells . . . 87

4.7 Combined Metformin and Ascorbic Acid can increase GATA6 Expression in two additional Aggressive Cell Lines 88
4.8 Ascorbic Acid increases GATA6 in CAPAN1 Cells 90
4.9 Combined Treatment can increase TET2, but not GATA6 in Healthy Immortalized Ductal Cells 91
4.10 Potential off target effects of Metformin and AA targeting GATA6 . . 94

5.1 Orthotopic Mouse Model Used for *in vivo* Experiments 98
5.2 Characterization of Orthotopic Tumours 100
5.3 Metformin/AA Increases 5'hmC 102
5.4 Metformin/AA Increases GATA6 103
5.5 RNA-seq of PANC1 tumours 105
5.6 Metformin/AA Induces a Partial Switch in Subtype in PANC1 Cells . 106
5.7 Metformin/ Ascorbic Acid Suppresses Wnt and Stem Cell Signalling Pathways in Tumour Cells 107
5.8 Metformin/ Ascorbic Acid Suppresses Wnt7b in Stroma Cells 108
5.9 Metformin/ Ascorbic Acid Inhibits nuclear β-Catenin in Tumour Cells 110
5.10 Metformin and Ascorbic Acid Inhibit Tumour Proliferation 112

6.1 SMAD4 Correlates with TET2 118
6.2 SMAD4 Promotes TET2 119
6.3 SMAD4 Knockdown Decreases TET2 and 5'hmC 120
6.4 SMAD4 Binds to an Enhancer at the TET2 Gene 122
6.5 SMAD4 Correlates with Molecular Subtype 123
6.6 SMAD4 Regulates Molecular Subtype Markers 124
6.7 SMAD4 Regulates Molecular Subtype Gene Sets 126

7.1 Potential Mechanisms of TET2 Loss in PDAC 138
7.2 Model for GATA6 Loss in Squamous Tumours 141

A.1 Molecular Subtyping of the Oxford Cohort 143
A.2 mRNA Expression of Subtype Markers in PDAC Cell Lines 144
A.3 TET2 Mutations in PDAC 144
A.4 Ascorbic Acid Treatment Does not Alter HIF-1α *in vitro* 145

List of Tables

2.1 **Cell Lines** 30
2.2 **qPCR Primers** 30
2.3 **Plasmids** 31
2.4 **Buffers** 31
2.5 **Kits** 32
2.6 **Antibodies** 32

3.1 **Indicators of Survival in the Oxford Cohort** 56

5.1 **Orthotopic Mouse Tumours** 98

List of Symbols and Acronyms

Abbreviations

$5'Aza$	5'Azacytidine
$5'caC$	5'carboxylcytosine
$5'fC$	5'formylcytosine
$5'hmC$	5'hydroxymethylcytosine/ Hydroxymethylation
$5'mC$	5'methylcytosine/ Methylation
α-KG	α-ketoglutarate
AA	Ascorbic Acid
$ADEX$	Aberrantly Differentiated Endocrine Exocrine
AML	Acute Myeloid Leukemia
BER	Base Excision Repair
BH	Benjamini-Hochberg Adjustment for Multiple Comparisons
CAP	Capecitabine
$CCLE$	Cancer Cell Line Encyclopedia
CpG	Cytosine-phosphate-Guanine
$CRUK$	Cancer Research UK
CSC	Cancer Stem Cell
DE	Definitive Endoderm
DFS	Disease-Free Survival

DIP	DNA Immunoprecipitation
$DMhP$	Differentially Hydroxymethylated Probe
DMP	Differentially Methylated Probe
$DMSO$	Dimethyl Sulfoxide
$DNMT$	DNA Methyltransferase
ECM	Extracellular Matrix
EMT	Epithelial to Mesenchymal Transition
ES	Embryonic Stem (Cells)
FDA	Food and Drug Administration
$FFPE$	Formal Fixed, Paraffin Embedded
FG	Foregut
$gDNA$	Genomic DNA
GEM	Gemcitabine
GEO	Gene Expression Omnibus
$GSEA$	Gene Set Enrichment Analysis
GT	Gut Tube
$H2O2$	Hydrogen Peroxide
$H\&E$	Haemotoxylin and eosin
$HSPC$	Hematopoetic Stem and Progenitor Cell
IGV	Integrated Genomics Viewer
IHC	Immunohistochemistry
IP	Intraperitoneal Injection
IPS	Induced Pluripotent Stem
$KSIC$	KRAS, SMAD4 and INK4A Mutant Mice (KRAS, SMAD4, INK4A, Cre)

MDS	Myelodysplastic Syndrome
MRI	Magnetic Resonance Imaging
OS	Overall Survival
$PAAD$	TCGA Pancreatic Adenocarcinoma Dataset
$PanIN$	Pancreatic Intraepithelial Neoplasia
PBS	Phospho-Buffered Saline
$PDAC$	Pancreatic Ductal Carcinoma
PE	Pancreatic Epithelial (Cells)
QC	Quality Control
$QM-PDA$	Quasimesenchymal
$TCGA$	The Cancer Genome Atlas
TF	Transcription Factor
TSG	Tumour Suppressor Gene
WT	Wild Type

Gene Symbols

$AMPK$	AMP-Activated Protein Kinase
APC	Adenomatous Polyposis Coli
$ARL6$	ADP Ribosylation Factor Like GTPase 6
$CA9$	Carbonic Anhydrase
$CDKN2A$	Cyclin Dependent Kinase Inhibitor 2A
$DAPK$	Death Associated Protein Kinase
EGF	Epidermal Growth Factor
$FIH-1$	Factor Inhibiting HIF-1
$GATA6$	GATA Binding protein 6

$GRB10$	Growth Factor Receptor Bound Protein 10
$HDAC$	Histone De-acetylase
HIF	Hypoxia-Inducible Factor
$HNF1B$	HNF1 Homeobox B
$KDM6A$	Lysine Demethylate 6A
KDR	Kinase Insert Domain Receptor
$KLF5$	Kruppel Like Factor 5
$KRAS$	KRAS proto-oncogene GTPase
$MBD2$	Methyl-CpG Binding Domain Protein 2
MML	Mixed Linease Leukemia 1
$MNX1$	Motor Neuron and Pancreas Homeobox 1
$mTOR$	mammalian Target of Rapamycin
$Nodal$	Nodal Growth Differentiation Factor
$NR5A2$	Nuclear Receptor Subfamily 5 Group A Member 2
$p300$	Histone Acetyltransferase p300
$PDX1$	Pancreatic And Duodenal Homeobox 1
PHD	Prolyl Hydroxylase Domain-Containing Protein
PRC	Polycomb Repressor Complex
$PRDM9$	PR/SET Domain 9
$RB1$	RB Transcriptional Corepressor 1
$RSPO1$	R-Spondin 1
$RUNX1$	RUNX Family Transcription Factor 1
Shh	Sonic Hedgehog
$SMAD4$	SMAD Family Member 4

*SOX*4	SRY-Box 4
TET	Ten-Eleven Translocation
TGD	Thymine DNA Glycosylase
*TP*53	Tumour Protein p53
*TP*63	Tumour Protein p63

Chapter 1

Introduction

1.1 Cancer

Cancer is a heterogeneous collection of diseases characterized by the uncontrollable growth of cells. As cancer cells progress, they evolve and become better able to evade the intracellular and extracellular factors that control normal homeostasis, allowing them to disseminate throughout the body and eventually leading to the death of their host. A number of changes commonly occur in most tumour types, referred to as the "Hallmarks of Cancer", see Fig. 1.1 (Hanahan and Weinberg, 2011). Central to these is the capability of tumour cells to proliferate in the absence of mitogenic signals. This typically occurs through somatic mutations in proto-oncogenes by genetic amplifications, activating mutations or removal of inhibitory domains. However, cancer signalling pathways are tightly regulated, and contain complex negative feedback loops in order to prevent any one signal from over stimulating cells. Sustained proliferation results in the release of anti-proliferative signals, as well as signals that tell cancer cells to exit the cell cycle or undergo apoptosis by activation of tumour suppressor genes (TSGs), which can be overcome by genomic deletions or inactivation mutations, as well as epigenetic mechanisms such as DNA methylation. Thus, cancer

is a multifactorial disease and multiple aberrations must accumulate in a cell before it can transform and become neoplastic. In order to grow beyond a certain size, tumours must become replicatively immortal, and promote the formation of blood vessels to supply them with oxygen and nutrients. Additionally, the immune system can recognize and eliminate cancer cells, and so additional mechanisms of immune evasion must arise for a tumour to form. Therefore, there are many barriers to tumorigenesis that a cancer cell must overcome. Approximately 37 trillion cells exist at any one time in the body, many of which are being continually replaced over a person's lifespan, yet only 50% of people will develop a cancer in their lifetime. It is likely that cancer is relatively rare in the body due to the barriers to tumorigenesis. However, tumour cells are in a constant state of evolution to bypass these mechanisms, and thus any tumours that do eventually arise are likely to be highly complex.



Figure 1.1: **The Hallmarks of Cancer**. A number of unifying traits underlie the development of cancer such as the ability to proliferate in the absence of mitogenic signals. Adapted from Hanahan and Weinberg (2011). Reprinted with permission of Elsevier with licence number 4674820712229.

According to Cancer Research UK (CRUK), cancer is the cause of over a quarter of all deaths in the UK. However, through an increased understanding of how tumour cells develop and progress, patient survival has dramatically improved, and in the

past 50 years survival rates have more than doubled from 25 to 50%.

1.1.1 Cancer Stem Cells

Another common feature of cancer cells is the loss of proper differentiation and reversion to a more stem-like state. Cancer cells and stem cells share many common features, such as the ability to indefinitely self renew. In many adult tissues, such as the intestine and epidermis, adult stem cells exist in dedicated niches where they can repopulate the surrounding tissue during normal homeostasis or injury (Clayton et al., 2007)(Barker et al., 2007). Interestingly, it is the stem cell niche, rather than the stem cells themselves, which is responsible for maintaining cells in a multipotent state. Upon leaving the niche daughter cells begin to differentiate, resulting in distinct hierarchies of cells within the tissue. Furthermore, if differentiated daughter cells re-enter the niche they can de-differentiate to become stem cells again (Tian et al., 2011)(van Es et al., 2012). The plasticity of differentiated cells to revert back to tissue specific progenitors, as well as the observation that most tumours are heterogeneous with cell hierarchies that frequently mirror their tissue of origin led to the hypothesis that stem or stem-like cells within the tumour are responsible for fueling most of the tumours growth. This is referred to the cancer stem cell hypothesis (Reviewed by Batlle and Clevers (2017)). One major implication of this theory is that that a single cancer stem cell (CSC) can repopulate an entire tumour in the same way that adult stem cells regenerate a tissue after injury. This, combined with the fact that cancer stem cells are resistant to conventional chemotherapy and radiation treatment may explain why most patients relapse after treatment.

If cancer stem cells are the drivers of tumorigenesis, then targeting these cells, or the niches that maintain them, may prevent tumour relapse following treatment. The cancer stem cell niche has successfully been targeted in some tumours. Colon stem cells are reliant on secreted Wnt and Epidermal Growth Factor (EGF) signals, and

CSCs in some colon tumours are still reliant on these signals, making them sensitive to EGF and Wnt inhibitors (Batlle and Clevers, 2017). However, many of the major driver mutations in cancer allow cells to become independent of the stem cell niche. For example, Adenomatous Polyposis Coli (APC) mutations in colorectal tumours activate Wnt signalling, while KRAS proto-oncogene GTPase (KRAS) mutations bypass EGF requirements (Fumagalli et al., 2017)(Fujii et al., 2016). Thus, as tumours develop they become more and more stem-like as they progressively build their own internal stem cell niche (Driessens et al., 2012).

1.1.2 Pancreatic Ductal Carcinoma (PDAC)

While survival rates have improved for almost all cancer types over the past 50 years, pancreatic ductal carcinoma (PDAC) continues to have a dismal outcome. Symptoms in most patients are vague, such as back pain or tiredness, meaning that the majority of tumours are locally advanced or metastatic by the time they get diagnosed. Additionally, tumours are recalcitrant to conventional therapy (resection and/or systemic Gemcitabine therapy), resulting in high rates of relapse. The combination of these factors means that only 2.5% of PDAC patients survive 5 years after diagnosis (CRUK).

In vertebrates the pancreas develops from ventral and dorsal domains of the gut endoderm epithelium. Around mouse embryonic day 8.5 the transcription factor Pancreatic And Duodenal Homeobox 1 (PDX1) can be detected in a subset of cells which become committed to a pancreatic fate and form the dorsal and ventral pancreatic buds. These buds then undergo a complex programme of proliferation, branching and differentiation in order to form the three compartments of the adult pancreas (Pictet et al., 1972)(Jonsson et al., 1994). The endocrine compartment of the pancreas, also known as the islets of Langerhans, are groups of cells that regulate glucose homeostasis by secreting hormones such as glucagon and insulin into the blood stream. The

exocrine compartment is composed of acinar cells that secrete digestive enzymes, and a ductal system that delivers these enzymes to the duodenum.

Over 90% of tumours that arise from the exocrine pancreas are PDAC (Kloppel and Luttges, 2001) and can develop from both ductal cells and acinar cells that undergo an acinar-to-ductal metaplasia as part of a wound repair process (Ferreira et al., 2017). Other pancreatic tumours can develop from endocrine cells (e.g. neuroendocrine tumours), but these are far rarer and have a better outcome than PDAC. An initial driver mutation is thought to cause the formation of a low grade pancreatic intraepithelial neoplasia (PanIN). In most cases this is due to a missense mutation at the guanine at position 12 (G12) of the KRAS protein, resulting in its constitutive activation. Additional deregulation of oncogenes and tumour suppressors causes further disruption of the normal tissue architecture, causing nuclei to become depolarized, and clusters of cells to bud away from the epithelium, eventually resulting in formation of a PDAC tumour.

While PDAC tumours share the hallmarks of many other cancer types, there are a few aspects that make PDAC tumours distinct. Firstly, PDAC tumours have a lower mutational burden than most other types of cancer (Humphris et al., 2017). Driver mutations in only four genes dominate: KRAS mutations are found in ∼90% of tumours, while Tumour protein p53 (TP53) mutations are present in over 70% of tumours and SMAD family member 4 (SMAD4) and Cyclin Dependent Kinase Inhibitor 2A (CDKN2A) mutations are both found in approximately 50% of patients (Smit et al., 1988)(Aguirre et al., 2018). A handful of other genes have been identified, but mutation rates in these are relatively low. Strikingly, there is little genetic heterogeneity even when comparing primary tumours to distant metastasis, suggesting that late stage PDAC is driven in a non-genetic manner (Reiter et al., 2018). Additionally, PDAC cells are typically accompanied by a dense stromal reaction, to the extent that neoplastic cells make up the minority of cells within the tumour. During

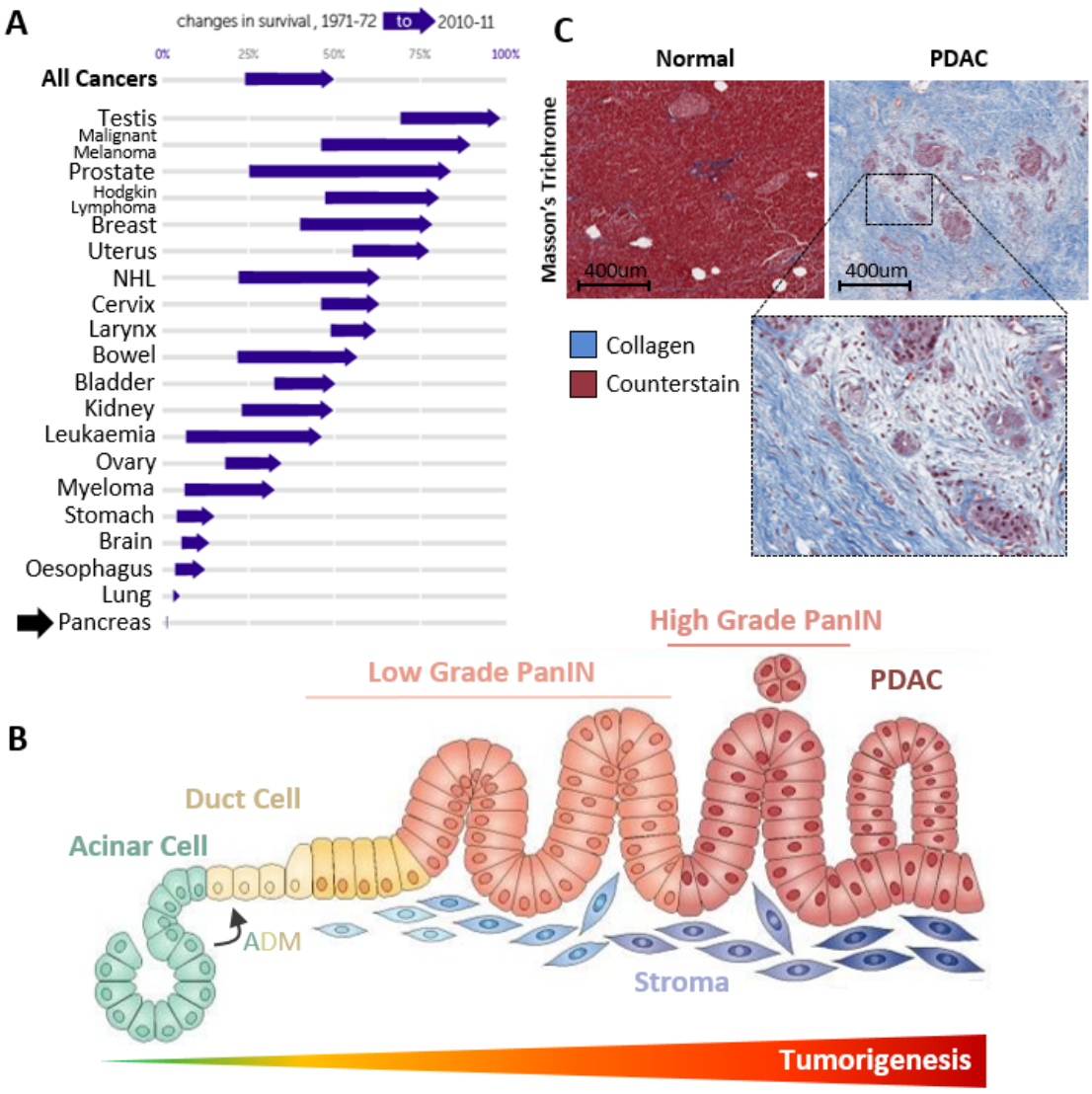


Figure 1.2: **Pancreatic Ductal Carcinoma (PDAC)**. **A**) Changes in overall survival for different cancer types between 1971 and 2011 (CRUK). Pancreatic cancer has the worst survival of all cancers, and has had almost no improvement in survival in the past 50 years. **B**) Formation and progression of PDAC. PDAC originates from acinar or ductal cells. Tissue architecture becomes more disrupted as PanINs form and develop into PDAC. Adapted from Morris et al. (2010). Reprinted with permission from Springer Nature with licence number 4674821208043. **C**) Example of (left) normal human pancreatic tissue and (right) a PDAC tumour stained with Masson's Trichrome. Collagen fibres are stained blue and the tissue is counter stained purple. Most of the tumour is composed of stromal cells which secrete collagen and growth factors.

PDAC formation, tumour cells release mitogenic substances that stimulate the invasion and proliferation of the surrounding stellate and fibroblast cells (Rasheed et al., 2012). In return, stromal cells secrete extracellular matrix (ECM) proteins such as collagens and fibronectin, as well as growth factors that support the growth of the tumour (Maitra and Hruban, 2008)(Thomas and Radhakrishnan, 2019). The stroma can also act as a cancer stem cell niche by secreting paracrine ligands such as Wnts and Nodal Growth Differentiation Factor (Nodal) that enhance the self renewal and invasiveness of the tumour cells (Lonardo et al., 2012)(Seino et al., 2018). Fig 1.2B shows an example of a human PDAC tumour stained with Masson's Trichome, which stains collagen fibres in the stroma blue (Stained by Zenobia D'Costa from the lab of Emmanouil Fokas). Small islands of tumour cells can be seen within a dense sea of stroma, constantly being supplied with growth factors that fuel their growth. The tumour microenvironment creates additional complications when attempting to treat PDAC, as the stroma can increase resistance to therapy by metabolizing gemcitabine and acting as a physical barrier for entry of small molecules to the tumour cells (Liang et al., 2017)(Provenzano et al., 2012).

1.1.3 Genomic Analyses of Cancer

Advances in genomic analyses and sequencing technologies have provided overviews of the genome, epigenome and transcriptome in unprecedented detail. Additionally, the dramatic decrease in the cost of these techniques means that they are now applicable to large cohorts of patients. The human genome project was the first attempt to sequence the genome, and took 13 years and $3-5 billion to complete. Now, whole genome sequencing can be performed for approximately $1000 in the space of weeks, leading to the creation of large patient cohorts of genome wide molecular databases such as The Cancer Genome Atlas (TCGA). This has helped to elucidate the molecular mechanisms that underlie pan-cancer tumour progression (Sanchez-Vega et al.,

2018)(Ding et al., 2018), and has also highlighted the differences between tumours of the same type.

Traditionally, tumours have been grouped by their tissue of origin, which has provided the basis for treatment options for each patient. However, better understanding of these tumours at the molecular level has shown that even morphologically similar tumours from the same cell of origin can develop into multiple different subtypes. Integrated clustering of genetic, epigenetic, proteomic and transcriptomic data from over 10,000 tumours covering 33 different cancer types found that, while the cell of origin was a dominant feature in determining the molecular subtype of a tumour, tumours arising from the same tissue could be highly heterogeneous (Hoadley et al., 2018). These subtypes express mutually exclusive sets of genes that can confer different resistances, vulnerabilities and clinical outcomes. Thus, tailoring therapy to each of these molecular subtypes, rather than simply the tissue of origin, may present the next stage of clinical diagnosis for targeted treatment of cancer.

1.1.4 Molecular Subtypes of PDAC

There have been three major studies into the molecular subtypes of pancreatic cancer using transcriptomic datasets. Initially, Collisson et al. (2011) described three subtypes; Classical, Quasimesenchymal (QM PDA) and Exocrine-like. QM PDA tumours were highly aggressive and were associated with a reduced dependence on KRAS signalling and lower expression of the endodermal differentiation marker GATA Binding Protein 6 (GATA6) relative to Classical tumours. Additionally, classical tumours had higher expression of epithelial genes such as E-Cadherin, while QM PDA tumours had high expression of mesenchymal associated genes such as N-Cadherin. Exocrine-like tumours on the other hand expressed genes involved in terminal pancreatic differentiation. However, classification of PDAC cell lines using the same gene signatures showed only two subtypes; QM PDA and Classical. This combined with its associa-

tion with terminal pancreatic differentiation suggests that the Exocrine-like subtype may be a result of contaminating exocrine tissue.

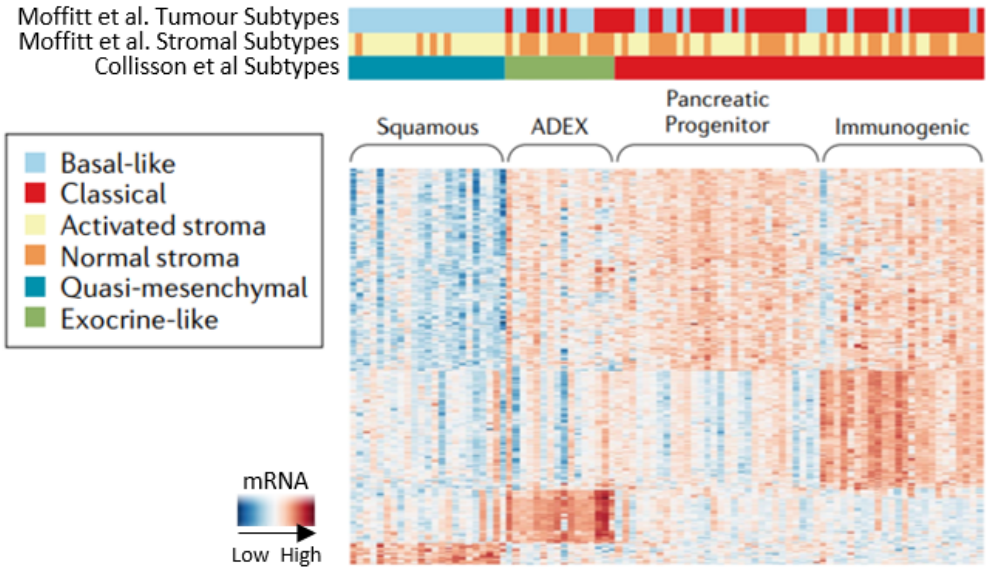


Figure 1.3: **Genomic Analyses of PDAC**. mRNA expression dataset classified according to Bailey et al. (2016) showing four groups of tumours that express different gene sets. Classification of the same tumours according to Moffitt and Collisson gene sets is shown in the panel above. Note that there is extensive overlap between aggressive molecular subtypes. Adapted from Collisson et al. (2019). Reprinted with permission from Springer Nature under licence number 4674821325001.

The idea that only two subtypes of PDAC exist was supported by Moffitt et al. (2015). Using transcriptomic data from differentiated tissue, Moffitt et al could exclude transcripts thought to be from normal pancreatic and stromal cells. Using this virtual microdissection technique, they identified two subtypes which they named basal and classical, as well as two stromal subtypes termed normal and activated that were independent of the subtypes of tumour cells. This classical subtype overlapped with the classical subtype described by Collisson and was also associated with increased GATA6 signalling, as well as higher SMAD4 mRNA. The exocrine-like gene signature on the other hand was found to be indistinguishable from healthy adjacent pancreatic tissue. The basal subtype partially overlapped with the QM PDA signature and was similarly associated with poorer patient survival.

Finally, Bailey et al. (2016) described four molecular subtypes: squamous, pancreatic progenitor, immunogenic and aberrantly differentiated endocrine exocrine (ADEX). The squamous subtype overlapped with QM-PDA tumours and partially overlapped with basal tumours. Again this subtype was associated with loss of GATA6 expression, as well as other endodermal differentiation markers such as PDX1, Motor Neuron and Pancreas Homeobox 1 (MNX1) and HNF1 Homeobox B (HNF1B) by promoter methylation. Disrupted epigenetic signalling within squamous tumours was also hinted at due to an increase prevalence for mutations in Lysine Demethylase 6A (KDM6A). Additionally, these tumours were enriched for TP53 mutations, activated EGF signalling and high expression of the oncogenic ΔN isoform of Tumour Protein 63 (TP63). ADEX tumours were similar to Exocrine-like in that they were enriched for terminal pancreatic differentiation genes, while progenitor tumours had characteristics similar to both classical subtypes described by Moffit and Collisson. Immunogenic tumours shared many characteristics of the progenitor subtype but with increased immune infiltration, making them potential targets for immunotherapy.

A number of features are common across these studies. Indeed a review authored by Collisson, Bailey and colleagues (Collisson et al., 2019) attempted to reconcile these results. QM PDA, basal and squamous are clearly distinct from the other molecular subtypes in their highly aggressive nature and loss of endodermal differentiation. The combination of these aggressive subtypes was referred to as squamous due to the similarity with other subtypes of the same name present in other types of tumours (Hoadley et al., 2014). The remaining subtypes were grouped together and referred to collectively as classical-pancreatic. These were characterized by being less aggressive than squamous tumours with higher expression of endodermal differentiation markers such as GATA6 and PDX1. Classical-pancreatic tumours could then be subdivided into classical-progenitor and ADEX. Classical-progenitor tumours consisted of the classical and progenitor subtypes, as well as immunogenic, which appear to be pro-

genitor tumours with immune infiltrates. The ADEX group was composed of ADEX and exocrine-like subtypes, which are characterized by high levels of terminal pancreatic differentiation genes which may or may not be due to contaminating healthy adjacent tissue. This squamous/classical-pancreatic classification is the nomenclature that will be used throughout this thesis.

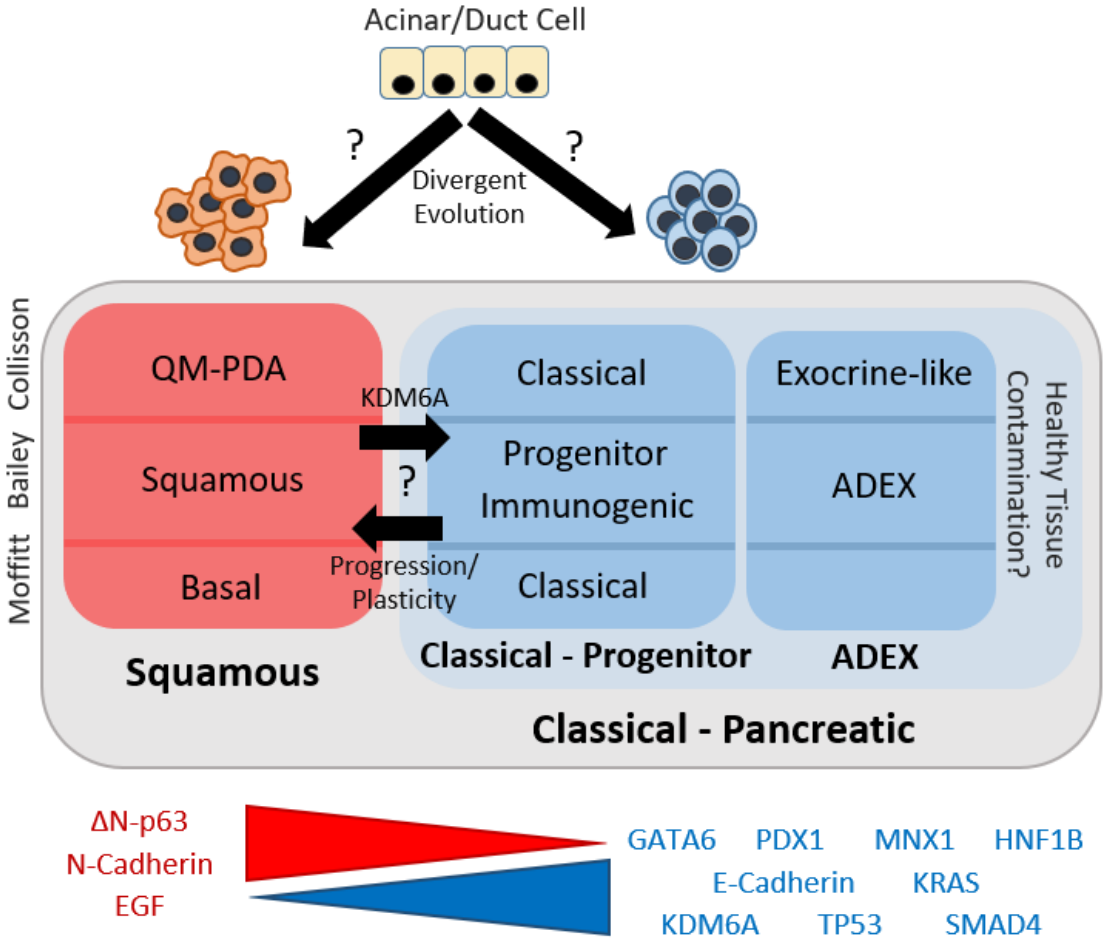


Figure 1.4: **Molecular Subtypes of PDAC**. Proposed nomenclature for molecular subtypes of PDAC. Collisson et al. (2019) reconciled previous studies to propose two subtypes; squamous and classical-pancreatic. Classical-pancreatic tumours can be further subdivided into classical-progenitor and ADEX. Subtypes may arise due to detergent evolution, or from progression of one subtype to another. Genes associated with a particular subtype are shown in the panel below.

1.1.5 Evolution of PDAC Molecular Subtypes

While the distinction between squamous and classical-pancreatic tumours has important implications for patient outcomes and potential treatment options, how these subtypes arise remains unclear. Evolutionarily, two possible models could explain the formation of two separate subtypes: a divergent evolution model whereby tumours develop as either squamous or classical-pancreatic depending on tumour initiation conditions, e.g. different driver mutations or cell of origin. This would suggest that each subtype is fixed and remains the same throughout the life of the tumour. Alternatively, subtypes may be plastic and progress from one to another as tumours become more aggressive.

The paucity of PDAC driver mutations makes it unlikely that that is a determining factor in subtype identity. TP53 and KDM6A mutations are enriched in squamous tumours. However TP53 is the second most common PDAC mutation and is also present in the majority of classical-pancreatic tumours. KDM6A mutations in contrast are only present in 6% of the PAAD TCGA cohort, and thus could only explain a minority of squamous tumours. PDAC tumours can arise from both acinar and ductal cells, and ductal derived tumours have been shown to be more aggressive than acinar (Ferreira et al., 2017), however there have not yet been any studies linking the cell of origin to molecular subtypes.

Mutation of KDM6A in KRAS mutant mouse cell lines resulted in increased expression of genes linked to squamous subtypes, such as increased ΔN-p63 and expression of mesenchymal associated genes, suggesting plasticity in these signatures and thus the subtypes themselves. Additionally, immunotherapy targeting macrophages within squamous tumours was shown to cause transcriptional changes consistent with a switch to ADEX or immunogenic subtypes (Candido et al., 2018), indicating that it is possible to modulate subtypes pharmacologically. Thus, conversion of squamous tumours towards classical-pancreatic, or preventing classical-pancreatic tumours from

developing into squamous, represents a potential therapeutic option to increase patient outcome.

1.1.6 GATA6

GATA6 expression was found the be differentially expressed between subtypes in all three studies by Collisson, Moffitt and Bailey. Strikingly, RNA *in situ* hybridization of GATA6 alone was found to be a robust surrogate biomarker for differentiating classical-pancreatic from squamous tumours in a clinical trial of advanced PDAC patients (Aung et al., 2018), suggesting that GATA6 is either a master regulator in determining the subtypes of these tumours, or is being co-regulated with other genes that are involved in determining subtype.

The GATA family is a evolutionarily conserved set of transcription factors that are critical for normal development and differentiation for eukaryotic organisms (Reviewed by Viger et al. (2008)). Each GATA family member consists of a conserved DNA-binding domain that recognizes the consensus binding site A/T-GATA-A/G flanked by N- and C-terminal domains that mediate protein-protein interactions (Kuo et al., 1997)(Patient and McGhee, 2002). Six family members exist in vertebrates: GATA1-3 are required for differentiation of hematopoietic lineages (Weiss and Orkin, 1995) and may also be involved in brain, spinal cord and inner ear development (Nardelli et al., 1999)(Lillevali et al., 2004). GATA4/5/6 on the other hand regulate differentiation of mesoderm and endoderm derived cell types (Molkentin, 2000). Early mouse knockout studies revealed a role for GATA4 in formation of the heart tube (Kuo et al., 1997)(Molkentin et al., 1997), GATA5 for female genitourinary tract development (Molkentin et al., 2000) and GATA6 for differentiation of the extraembryonic endoderm (Morrisey et al., 1998)(Koutsourakis et al., 1999).

GATA4 and GATA6 have overlapping expression profiles in the early pancreas, but in the adult pancreas GATA4 becomes localized to the exocrine compartment, while

GATA6 is expressed in both exocrine and endocrine tissue (Decker et al., 2006)(Ketola et al., 2004)(Sartori et al., 2014). Genome sequencing of neonatal diabetic patients with defective pancreatic development found that over half of cases had spontaneous GATA6 somatic mutations (Allen et al., 2011). Additionally, homozygous mutation of GATA6 in patient derived induced pluripotent stem (IPS) cells completely abrogated definitive endoderm differentiation due to loss of the endodermal transcriptional programme, while heterozygous mutations impaired pancreatic lineage commitment and halved the number of $PDX1^{+}$ pancreatic progenitors (Chia et al., 2019)(Tiyaboonchai et al., 2017). This may be due to a loss of PDX1 expression as GATA6 can directly bind to GATA motifs within the PDX1 gene (Carrasco et al., 2012). Thus in humans GATA6 is a critical regulator of pancreatic specification and differentiation. However in mice there appears to be functional redundancy between GATA6 and GATA4 as mutation of either gene alone does not affect pancreatic development, while double mutation resulted in impaired growth of the pancreatic epithelium, reduced proliferation of pancreatic cells and loss of endocrine and exocrine lineage commitment (Carrasco et al., 2012).

Initial studies of GATA6 in PDAC revealed a subset of tumours with GATA6 amplifications. This combined with the fact that overexpression of GATA6 in PDAC cell lines increased proliferation rates suggested that GATA6 had an oncogenic function (Kwei et al., 2008)(Fu et al., 2008). However, patients with GATA6 amplifications typically had well differentiated tumours that responded better the therapy (Zhong et al., 2011). Later studies showed that GATA6 inhibits epithelial to mesenchymal transition (EMT) by regulating core EMT transcription factors, and can prevent tumour dissemination *in vivo*, indicating that GATA6 has tumour suppressive as well as oncogenic roles (Martinelli et al., 2017). However, amplification of GATA6 occurs during PanIN formation (Zhong et al., 2011), indicating that this is an early stage event in PDAC formation. Therefore, GATA6 may be upregulated initially and then

later lost as squamous tumours develop.

1.1.7 Canonical Wnt signalling and the PDAC Stem Cell Niche

The Wnt pathway is highly conserved signalling pathway that has been linked to stem cell maintenance and tumour progression in many tumour types (Reviewed by Ghosh et al. (2019)). In the absence of active canonical Wnt signalling, the transcription factor β-catenin is phosphorylated by the Axin complex, leading to its ubiquitination and degradation. A Wnt signalling cascade is activated in response to secreted Wnt ligands binding to receptors at the cell membrane that results in inhibition of the Axin complex. This allows β-catenin to accumulate and translocate to the nucleus, where it interacts with the TCF/LEF family of transcription factors to drive expression of its target genes (Reviewed by MacDonald et al. (2009)). The Wnt pathway has been extensively studied in the intestine as it is responsible for maintaining the intestinal stem cell niche. Here, paneth cells secrete Wnt ligands that activate Wnt signalling in columnar stem cells at the base of intestinal crypts. Cells that become independent of these secreted Wnts, e.g. through mutation of members of the Axin destruction complex, are often the founders of colorectal tumours (Zhang and Shay, 2017).

Canonical Wnt signalling is required for PDAC initiation as mutation of β-catenin or inhibition of Wnt ligand secretion prevents PanIN formation (Zhang et al., 2013). Additionally, enhanced Wnt signalling has been linked to decreased survival (Sano et al., 2016), and single cell RNA-seq of circulating PDAC tumour cells implicated Wnt signalling in the distant spread of PDAC tumours (Yu et al., 2012). A number of Wnt inhibitors have been tested in clinical trials for PDAC, and have been shown to promote tumour differentiation and decrease metastasis (Reviewed by Krishnamurthy and Kurzrock (2018)). However, while most PDAC tumours show elevated β-catenin, this is typically found the be independent of Wnt pathway mutations (Zeng et al.,

2006)(Seino et al., 2018).

The poorly differentiated phenotype of squamous tumours suggests that regulation of stemness is critical for PDAC progression. Wnt ligands are critical components for the continued proliferation of healthy pancreatic organoids (Broutier et al., 2016), suggesting that they are required for maintenance of the pancreatic stem cell niche (Murrow et al., 2017). Interestingly, Seino et al. (2018) found that organoids derived from PDAC tumours exhibit different Wnt niche requirements, with some being dependent on external Wnt ligands and others able to autonomously activate their own Wnt signalling. Wnt autonomous organoids were more aggressive and more squamous in identity, in line with the idea that Wnt signalling is driving PDAC tumour stemness. Strikingly, loss of GATA6 caused Wnt-dependent organoids to become Wnt-autonomous. Therefore, the loss of GATA6 seen in squamous tumours may function to drive tumour stemness by allowing tumour cells to produce their own Wnt signalling. This may allow tumour cells to colonize distant regions of the body as they are independent of stromal cell support.

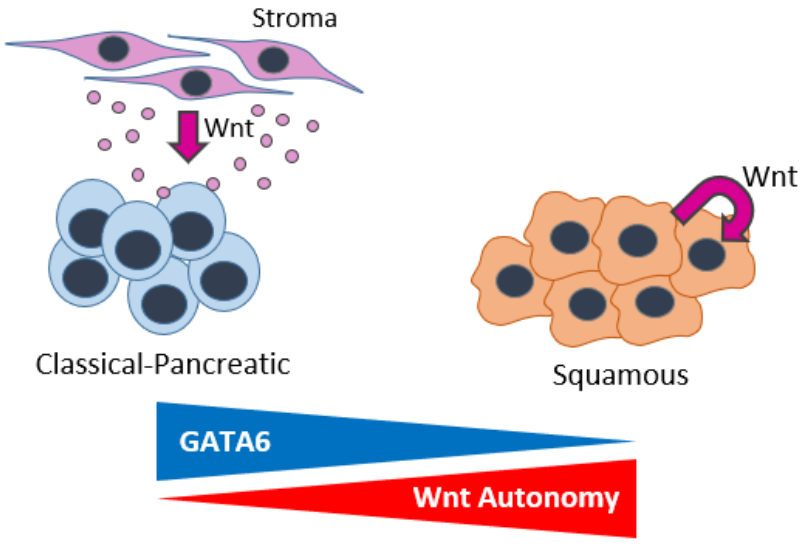


Figure 1.5: **Loss of GATA6 Promotes Wnt Autonomy and Squamous Subtypes** Loss of GATA6 allows tumours to become independent of the surrounding stroma and activate their own Wnt signalling leading to development of a squamous tumours.

From the above data a hypothesis for how a squamous subtype is acquired can be drawn: An acinar or ductal cell gives rise to a precursor PanIN lesion that progresses to a PDAC tumour that is dependent on the surrounding stroma to supply it with Wnt ligands. At some point GATA6 expression is epigenetically lost. This may occur as the tumour is initially forming, resulting in the immediate development of a squamous tumour, or in an established classical-pancreatic tumour which then switches subtype to become squamous. Loss of GATA6 results in downregulation of the classical-pancreatic gene expression profile, and upregulation of squamous genes and enables the tumour to activate its own Wnt signalling. With the tumour independent of the stroma the tumour becomes more aggressive and better able to metastasize (Fig. 1.5). Key to this process is the loss of GATA6, the cause of which remains unclear.

1.2 Epigenetics

In 1956, Conrad Waddington successfully influenced the phenotype of an organism using an environmental stimulus (C.H, 1956). Fruit fly embryos exposed to ether gas or heat shocks developed a bithorax phenotype that could be inherited over multiple generations. These changes occurred to the whole group of flies over a short time frame of a few generations, making it unlikely that the phenotypes seen were genetically driven. Thus, the term "epigenetics" was coined to describe heritable phenotypic changes that occur without a change in genotype. In 1975, an adult frog skin cell was transplanted into an enucleated donor egg, leading to the development of a cloned frog and showing that differentiated cells maintain a full complement of genes and do not lose any as they differentiate (Gurdon et al., 1975). Therefore, the remarkable range of phenotypic differences seen between somatic cells in multicellular organisms is a consequences of their epigenetics, and not their genetics.

While genetic mutations are considered the main drivers of tumorigenesis, it is becoming increasingly evident that epigenetic changes play a key role in tumour initiation, progression and metastasis (Chatterjee et al., 2018). This is particularly true for tumours with low mutational burdens such as PDAC. Genetically, there is little difference between primary PDAC tumours and distant metastasis (Reiter et al., 2018), but metastatic tumours display widespread epigenetic reprogramming (McDonald et al., 2017). Similarly, very few mutations have been linked to squamous tumours, with one of the only exceptions being an epigenetic enzyme (KDM6A), while loss of endodermal differentiation in squamous tumours is epigenetically driven (Bailey et al., 2016). Therefore, in PDAC, aggressive phenotypes and molecular subtypes may be epigenetically driven. Critically, epigenetic mechanisms are reversible, making them easier to target therapeutically than most genetic changes.

Today, epigenetics refers to heritable changes in gene expression that occur without a change in the underlying DNA sequence. This applies to any change that

persists for more than a single cell replication and can include covalent modifications to the DNA, changes in the way DNA is packaged, modifications to RNA and certain transcription factors that maintain their own expression.

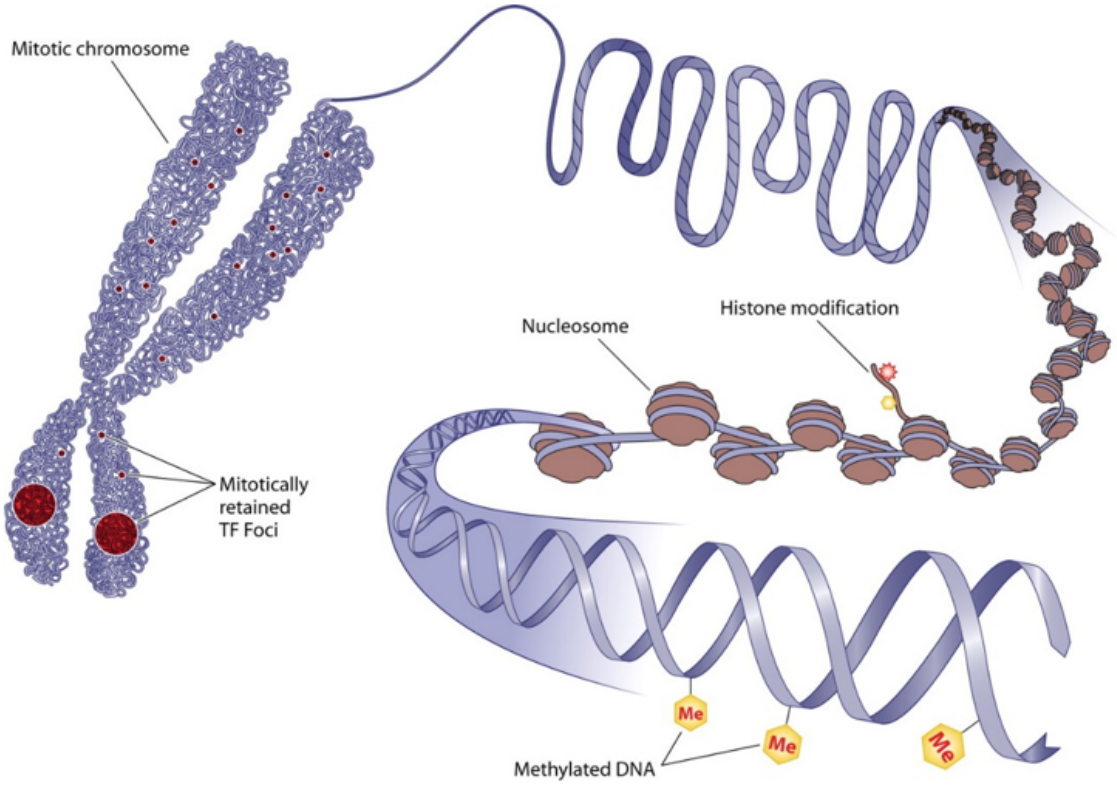


Figure 1.6: **Epigenetic Mechanisms** Histone modifications and DNA methylation are two examples of epigenetic mechanisms that can alter gene expression without affecting the underlying DNA sequence. Adapted from Zaidi et al. (2010a). Reprinted with permission from the American Society for Microbiology.

1.2.1 Histones

Histone modifications are some of the best characterized epigenetic mechanisms. DNA is not free floating within the nucleus but is instead packaged as chromatin, the basic building block of which is the nucleosome. This allows the approximately 2 metres of DNA within a human cell to fit into a space only a few micrometres across. 127 base pairs of DNA wrap around each nucleosome, which is composed of an octamer

of four core histone proteins: H2A, H2B, H3 and H4. Each of these histones has a tail that contains multiple residues that can have covalent modifications such as acetyl, methyl or phosphate groups attached (Fischle, 2005). These modifications can then act as recognition sites to recruit proteins that alter the chromatin structure and accessibility of DNA (Jenuwein and Allis, 2001).

Each nucleosome can contain a variety of histone modifications, and it is the combination of modifications, referred to as the "histone code", that determines the accessibility and activity of the underlying DNA. Histone enzymes can either establish new modifications (writers), or remove existing marks (erasers). The development of CRISPR/Cas9 based technology has helped to clarify many aspects of the histone code by targeting the catalytic domains of histone enzymes to gene regions to determine their direct effects. For example, PR/SET domain 9 (PRDM9) induces trimethylation of lysine 4 of histone 3 (H3K4me3) and is a hallmark of gene expression and is present at most active promoters (Dong et al., 2012). Targeting the catalytic domain of PRMD9 to a promoter region is sufficient to induce transient gene expression, but sustained gene activation is only achieved when H3K4me3 was induced in combination with trimethylation of histone 3 lysine 79 (H3K79me3) (Cano-Rodriguez et al., 2016).

1.2.2 DNA Methylation

DNA methylation is one of the first epigenetic marks shown to be altered in cancers, with global methylation changes being reported in tumours relative to healthy tissues in 1983 (Feinberg and Vogelstein, 1983). Since then aberrant DNA methylation has been shown to be a pan-cancer phenomenon (Saghafinia et al., 2018). Methylation occurs at Cytosine-phosphate-Guanine (CpG) dinucleotides by attachment of a methyl group (CH2) at position 5-carbon position of the cytosine residue (5'methylcytosine, 5'mC) by DNA methyltransferase enzymes (DNMTs) (Bestor and Ingram,

1983). Regulatory regions such as promoters and enhancers often contain clusters of CpG dinucleotides, referred to as CpG islands. 5'mC can silence these regions by multiple mechanisms, including obscuring transcription factor binding sites, creating new binding sites for silencing transcription factors, and recruiting repressive histone enzymes via Methyl-CpG-binding domain proteins (MBD1-3) (Curradi et al., 2002). 60-80% of CpG sites within the mammalian genome are modified by 5'mC, and the dynamic regulation of these sites is essential for transcriptional regulation during development (Smith and Meissner, 2013)(De Carvalho et al., 2010)(Meissner, 2010).

Tumour suppressor silencing by 5'mC is widespread during tumorigenesis (Chatterjee et al., 2018), leading to the search for DNMT inhibitors in an attempt to reactivate these genes. 5-azacytidine (5'Aza) is a cytosine analogue that incorporates into DNA and acts as a suicide substrate for DNMTs, irreversibly inhibiting their activity and leading to widespread de-methylation (Christman, 2002). 5'Aza treatment has been shown to induce TSG reactivation in multiple tumour types (Ghoshal et al., 2005)(Patra et al., 2011)(Corn et al., 2000) and was approved for treatment of myelodysplastic syndrome (MDS) by the Food and Drug Administration (FDA) in 2004 (Raj and Mufti, 2006). Furthermore, PanIN formation in mice with activating KRAS mutations (G12D) can be almost completely suppressed upon 5'Aza treatment (Grasso et al., 2015), implicating DNMT inhibition as a potential therapy for treating or preventing PDAC. However, 5'Aza can also covalently bind DNMTs to DNA, resulting in DNA damage and cytotoxicity that causes side effects in patients which has limited its use (Juttermann et al., 1994)(Yang et al., 2006)(Karahoca and Momparler, 2013).

1.2.3 de-Methylation and Hydroxymethylation (5'hmC)

There are two types of DNMTs in mammals: DNMT3A/B deposit *de novo* methylation, while DNMT1 maintains methylation following DNA replication by recognizing hemimethylated DNA and methylating the nascent DNA strand (Bostick et al., 2007)(Hermann et al., 2004). In the absence of DNMT1 5'mC becomes passively diluted with each replication cycle. This occurs in the maternal genome during preimplantation development through active exclusion of DNMT1 from the nucleus of early embryos in order to deplete 5'mC genome wide (Seisenberger et al., 2013). The paternal genome on the other hand loses 5'mC at a rate too rapid to be explained by passive de-methylation (Mayer et al., 2000), suggesting that active methods of de-methylation also exist.

It was only in 2009 that the DNA de-methylating enzymes were identified when Tahiliani et al. (2009) showed that Ten-Eleven Translocation 1 (TET1) is capable of oxidizing 5'mC to 5'hydroxymethylcytosine (Hydroxymethylation, 5'hmC). Later TET2 and TET3 were also shown to possess the same ability (Ito et al., 2010). 5'hmC can be further modified by TET proteins to 5'formylcytosine (5'fC) and 5'carboxylcytosine (5'caC). 5'fC and 5'caC are recognized by the DNA missmatch repair protein Thymine DNA Glycosylase (TDG) and excised from the genome, leaving an abasic site that can be restored to an unmodified cytosine through the base excision repair pathway (BER) (He et al., 2011)(Maiti and Drohat, 2011)(Weber et al., 2016). Thus, a methylated cytosine can be actively removed and replaced.

While TET enzymes recognize 5'mC, 5'hmC and 5'fC as substrates, the conversion of 5'mC to 5'hmC is 3-5 times faster than the other reactions (Ito et al., 2011)(Hashimoto et al., 2012)(Hu et al., 2015). 5'hmC can be stably maintained in the genome as it can accumulate to high levels in organs such as the brain (Kriaucionis and Heintz, 2009) and in mouse ESCs certain CpG sites are modified with 5'hmC and not 5'fC or 5'caC (Wu et al., 2011)(Xia et al., 2015) suggesting that 5'hmC

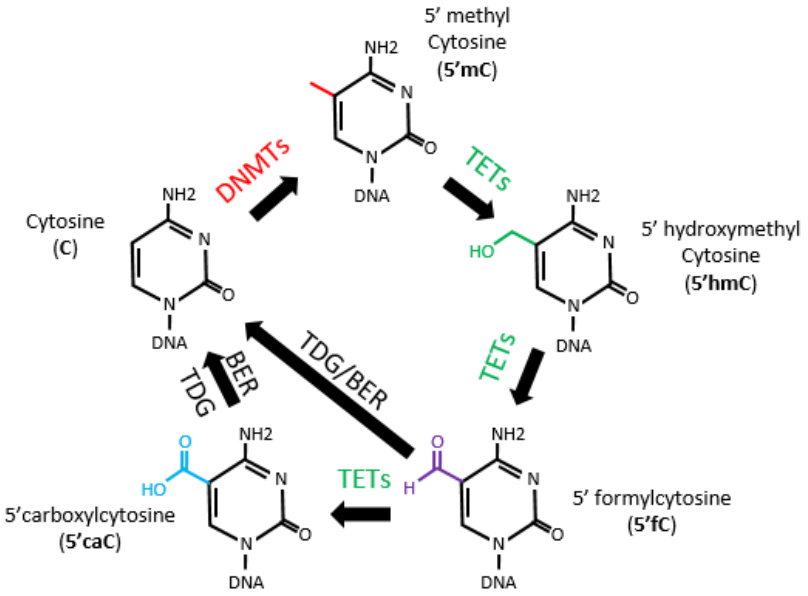


Figure 1.7: **Cytosine Modifications**. Schematic showing modifications to cytosines by DNMT and TET enzymes. Unmodified cytosines (C) can become methylated by DNMTs (5'mC). 5'mC is converted to 5'hmC, 5'fC and 5'caC by TET enzymes. 5'caC and 5'fC are recognized by TDG and the DNA damage repair pathway and removed by base excision repair (BER) and replaced by an unmodified cytosine.

may have other functions that just being an intermediate in the de-methylation pathway. Unlike 5'mC, 5'hmC is compatible with active transcription (Johnson et al., 2016). Additionally, as DNMTs cannot recognize 5'hmC as a substrate, the presence of 5'hmC can inhibit DNMT activity up to 60-fold (Hashimoto et al., 2012)(Inoue and Zhang, 2011)(Ji et al., 2014). Therefore, the presence of 5'hmC at a regulatory region is proposed to keep it in an active state by preventing its methylation.

1.2.4 5'hmC Detection Methods

Various methods for mapping the genomic distribution of 5'hmC have been developed (Reviewed here: Wu and Zhang (2017); Skvortsova et al. (2017)). Most of these methods can broadly be categorized into two techniques: antibody based and bisulfite based. Antibody based techniques involve DNA immunoprecipitation (DIP) using antibodies specific for 5'hmC (5'hmC-DIP), or the labelling of 5'hmC with molecules

such as biotin which are then immunoprecipitated (e.g. hME-Seal). These methods capture and enrich for DNA fragments containing 5'hmC but do not specify where in the fragment the modified base is and therefore do not provide base pair resolution. Furthermore, the relative enrichment of fragments is determined and not an absolute level of 5'hmC. Some bisulfite based techniques on the other hand provide base pair resolution of 5'hmC distribution at absolute levels. Under conventional bisulfite conditions 5'mC and 5'hmC are indistinguishable. Therefore, these techniques rely on altering 5'hmC prior to bisulfite conversion. This can be done directly by treating DNA with TET proteins to further oxidize 5'hmC (TAB-Seq) or by treating DNA with chemicals that selectively modify 5'hmC and not 5'mC (oxBS-Seq). However, the reagents used for these techniques are difficult to work with which has limited their use for more difficult types of samples, such as formalin fixed, paraffin embedded (FFPE) tissues.

1.2.5 5'hmC in the Pancreas

At the time of writing of this thesis, only a single major paper into the role of 5'hmC in the pancreas has been published. Li et al. (2018) measured changes in 5'hmC when embryonic stem (ES) cells were differentiated to pancreatic epithelial (PE) cells via the major endodermal intermediates: definitive endoderm (DE), gut tube (GT) and foregut (FG) cells. In ES cells 5'hmC is important for maintaining pluripotency as mutation of TET genes results in spontaneous differentiation and widespread CpG methylation (Tovy et al., 2017). Upon ES differentiation dynamic changes were seen in both the global levels of 5'hmC and distribution within the genome. 5'hmC was abundant in ES cells, while intermediate progenitors (DE, GT and FG) had 50-60% lower global levels. Differentiated pancreatic cells in contrast had similar levels of 5'hmC to ES cells. When the distribution of 5'hmC was analyzed, 5'hmC positively correlated with active regulatory regions such as promoters and enhancers at each

differentiation stage. Interestingly, 5'hmC peaks accumulated at either side of these regulatory regions, rather than within the region itself. These 5'hmC "rims" demarcated 5'mC free regions, consistent with the role of 5'hmC in antagonizing 5'mC. Critically, 5'hmC accumulated at key pancreatic differentiation genes such as PDX1 in PE cells, suggesting that 5'hmC is involved in the activation and/or maintenance of these genes. Thus, during pancreatic differentiation 5'hmC is globally lost and reestablished at pancreatic specific genes. Expression of the different TET genes was also dynamically regulated during this process. TET1 is highly expressed in ES cells, but becomes gradually depleted and replaced by TET2 and TET3 as cells differentiate through the intermediate progenitors. In differentiated pancreatic cells TET3 is lost and only TET2 remains, suggesting that TET2 is the main enzyme responsible for establishing the 5'hmC landscape in the adult pancreas (Li et al., 2018).

1.2.6 TETs

TET proteins are Fe^{2+} and α-ketoglutarate (α-KG) dependent dioxygenases. The C-terminal region of TETs consists of a double-stranded β-helical structure that brings together α-KG, Fe^{2+} and 5'mC for oxidation (Pastor et al., 2013)(Hu et al., 2013). Both TET1 and TET3 contain a CXXC domain which can recognize and bind to a single CpG site (Zhang et al., 2010)(Xu et al., 2011). TET2 on the other hand has no discernible CXXC domain, or other domains that can directly bind to DNA, and instead likely interacts with other DNA binding partners that allows for more dynamic, context dependent alterations in 5'hmC (Rasmussen and Helin, 2016). For example, WT1 can interact with and recruit TET2 to proliferation genes in order to increase their expression in leukemia cells (Wang et al., 2015), while in osteoclast cells TET2 interacts with RUNX Family Transcription Factor 1 (RUNX1) to promote the expression of differentiation associated genes (Chu et al., 2018). Therefore, TET2 function is often context specific and dependant on which binding partners are present

in each cell.

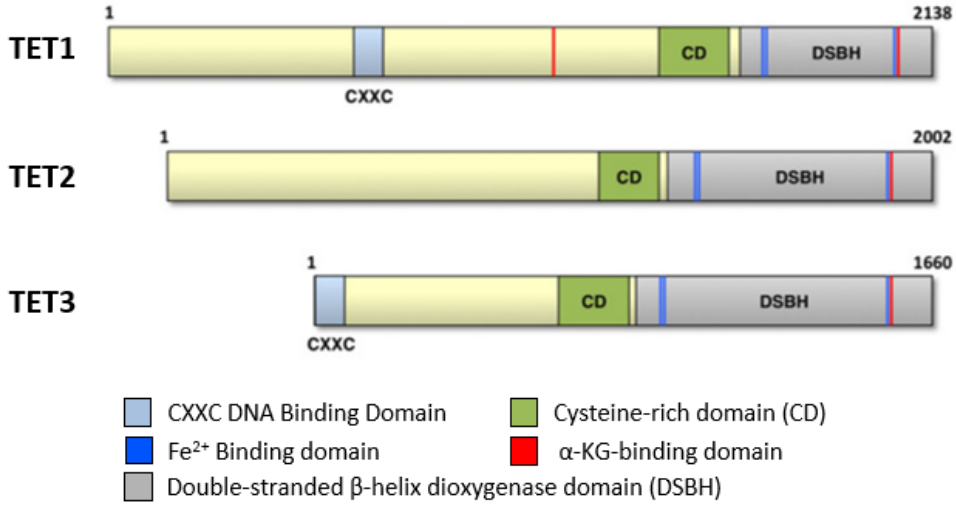


Figure 1.8: **TET Proteins** Structure of the TET family proteins. Adapted from Nakajima and Kunimoto (2014). Reprinted with permission from Wiley Publishing Asia.

Altered TET activity has been linked to cancer even before their role as 5'hmC enzymes was elucidated. The name Ten-Eleven Translocation was given to the TET family following the discovery of a translocation between the TET1 gene on chromosome 10 and Mixed Lineage Leukemia 1 (MLL1) gene on chromosome 11 in a small proportion of patients with acute myeloid leukemia (AML), referred to as a t(10;11) AML (Ono et al., 2002)(Lorsbach et al., 2003). Later deep sequencing of AML tumours showed that loss of function mutations in TET2 occur in 12-32% of patients, as well as a range of other hematological diseases (Weissmann et al., 2012)(Rasmussen and Helin, 2016)(Scourzic et al., 2015). Patients with these mutations had lower global levels of 5'hmC compared to TET2 wild type (WT) patients (Ko et al., 2010), as well as a worse prognosis (Chou et al., 2011). TET2 is highly expressed in hematopoetic cell lineages (Ko et al., 2010)(Moran-Crusio et al., 2011) and TET2 mutant mice displayed an expansion of the hematopoetic stem and progenitor cell (HSPC) population and disrupted HSPC differentiation (Ko et al., 2010), suggesting that TET2 is important for hematopoetic stem cell maintenance.

In many hierarchically organized tissues such as the skin, colon and pancreas global 5'hmC levels correlate with differentiation status, with the most differentiated cells having the highest levels of 5'hmC and progenitor cells having the lowest (Haffner et al., 2011). Additionally, global loss of 5'hmC relative to healthy tissue has been reported for a variety of solid tumours, including melanoma, liver, lung, breast and kidney cancer (Liu et al., 2013)(Yang et al., 2013)(Lian et al., 2012), suggesting that loss of 5'hmC may be involved in the de-differentiation of cancer cells. However, unlike in hematopoetic diseases, TET mutations in solid cancers are relatively rare, suggesting non-genomic methods of TET disruption in most cases (TCGA). Because of this, TET proteins are often still present at low levels, and by enhancing their activity through addition of co-factors the normal 5'hmC landscape can be restored, causing cells to differentiate (Rasmussen and Helin, 2016). In the pancreas $PDX1^{+ve}$ progenitor cells are marked by very low 5'hmC which progressively increases as cell differentiate (Gurlo T, 2018)(Oral Presentation). However there are very few studies on 5'hmC changes in PDAC, with only a single study in PDAC cell lines showing redistribution of 5'hmC relative to immortalized pancreatic cells (Bhattacharyya et al., 2017) and no studies using patient tumours. Critically, TET2 stability is regulated by a number of environmental factors that are drastically altered during PDAC development such as hypoxia via Hypoxia-inducible factor 1 alpha (HIF-1α) (Fischer and Miles, 2017) and glucose levels via AMP-activated protein kinase (AMPK) (Wu et al., 2018). Thus is it highly likely that TET2 will be disrupted as PDAC tumours develop.

1.3 Aims

Genomic analysis have identified highly aggressive subtypes of PDAC but it remains unknown how these subtypes arise and develop. Two things are clear: loss of endodermal differentiation is a key component of squamous subtypes, and this process is likely to be epigenetically driven. 5'hmC is an epigenetic mark that is dynamically regulated during pancreatic differentiation and accumulates at important differentiation genes but studies into 5'hmC in PDAC are lacking. Therefore, the aims of this thesis are to investigate the role of 5'hmC during PDAC progression; focusing on tumour de-differentiation and development of molecular subtypes.

Chapter 2

Materials and Methods

2.1 Materials

All chemicals were purchased from either Sigma-Aldrich or Thermo Fisher Scientific and used according to manufacturer's guidelines, unless stated otherwise. Tissue culture plastic was purchased from Corning and Sarstedt. Glassware was purchased from VWR and Fisher. Experiments are broadly split into 4 areas: generation of genomic information from a cohort of FFPE PDAC patient tissue, bioinformatics and statistical tests, molecular biology and *in vitro* experiments, and orthotopic *in vivo* experiments and histological analysis. Unless stated otherwise, all experiments were performed by the author.

2.2 Tables

Table 2.1: **Cell Lines**

Cell Line	Subtype	KRAS	TP53	SMAD4	Doubling Time
PANC1	Aggressive	G12A	273H	WT	40 Hours
PSN1	Aggressive	G12R	132Q	HD	20 Hours
MiaPaca2	Aggressive	G12C	248T	WT	40 Hours
CAPAN-1	Classical	G12V	159V	343-Stop/ 577L	50 Hours
DEC-hTERT	Immortalized Ductal Cells	WT	WT	WT	24 Hours

Pancreatic cell lines used for experiments. Molecular subtypes determined in Chapter 4, Figure 1 are shown, as well as approximate doubling time and mutation status of KRAS, TP53 and SMAD4. WT = Wild Type, HD = Homozygous Deletion. Mutational data was taken from the Catalogue Of Somatic Mutations In Cancer (COSMIC) database, as well as Moore et al. (2001) and Deer et al. (2010).

Table 2.2: **qPCR Primers**

Target	Sequence	PrimerBank id.
GAPDH	GGAGCGAGATCCCTCCAAAAT	378404907c1
	GGCTGTTGTCATACTTCTCATGG	
SMAD4	CTCATGTGATCTATGCCCGTC	195963400c1
	AGGTGATACAACTCGTTCGTAGT	
GATA6	CTCAGTTCCTACGCTTCGCAT	40288196c1
	GTCGAGGTCAGTGAACAGCA	

Forward and Reverse primer sequences used for qPCR

Table 2.3: **Plasmids**

Construct	**Backbone**	**Provider**
pcDNA3	pcDNA3	Addgene
FLAG-TET2	pcDNA3	Addgene #60939
FLAG-SMAD4	pcDNA3	Addgene #80888

Plasmids used for overexpression experiments

Table 2.4: **Buffers**

Buffer	**Contents**	**Protocol**
Phosphate-Buffered Saline (PBS)	137mM NaCL, 2.7mM KCL, 10mM Na2HPO4, 1.76 mM KH2PO4, ph7.4	Various
PBS Tween (PBS-T)	PBS as above, 0.1% (v/v) Tween 20	Various
Laemmli Lysis Buffer	50mM Tris-HCL pH6.8, 2% (w/v) SDS, 1X Protease Inhibitor cocktail, 1X Phosphatase inhibitor cocktail	Protein Isolation
SDS-Page Running Buffer	25mM Tris, 190mM Glycin, 0.1% (w/v) SDS	Western Blot
SDS-PAGE Sample Buffer	10% (v/v) Glycerol, 62.5 mM Tris-HCL ph6.8, 2% (w/v)SDS, 2.89mM Bromophenol Blue, 2% (v/v) 2-mercaptoethanol	Western Blot
Western Blot Transfer buffer	25mM Tris, 190mM Glycine, 20% (v/v) Methanol	Western Blot
5% Skimmed milk blocking buffer	5% (w/v) skimmed milk powder in 1X PBS-T	Western Blot, Dot Blot
DNA Lysis Buffer	100mM Tris-HCL (pH8.5), 5mM EDTA, 0.2% SDS, 100mM NaCL, 0.5mg/ml Proteinase K	DNA Isolation for Dot Blot
DNA Denaturing Buffer	200mM NaOH, 20mM EDTA	Dot Blot
Tris-EDTA Buffer (pH 9)	100mM Tris Base, 1mM EDTA, 0.05% Tween 20	IHC

Buffers used for experiments. Buffers used as part of a kit are not included.

Table 2.5: **Kits**

Kit	**Provider**	**Cat.no**	**Protocol**
RNeasy FFPE Kit	Qiagen	73504	RNA isolation from FFPE Scrolls
DNA FFPE Tissue Kit	Qiagen	56404	DNA isolation from FFPE Scrolls
AllPrep DNA/RNA/miRNA Universal Kit	Qiagen	80224	DNA, RNA and miRNA isolation from orthotopic tumours
Masson's Trichome Stain Kit	Sigma	HT15-1KT	Masson's Trichrome Stain
CEGX TrueMethyl Array Kit	CEGX	0414-32	Genome wide Methylation and Hydroxymethylation
Power SYBR Green Cells to CT Kit	Thermo Fisher	4402953	RNA isolation and cDNA synthesis from cell lines

Kits used for experiments. Standard protocols for each kit can be found on the providers website. Any deviations from the standard protocol are outlined in this chapter.

Table 2.6: **Antibodies**

Antibody	**Cat.no**	**WB Dilution**	**IHC Dilution**	**Provider**
5'hmC	39769	1:1000	1:100	Active Motif
Actin	sc-47778	1:5000		Santa Cruz
AMPK	2793	1:1000		Cell Signalling
pAMPK T172	2535S	1:500		Cell Signalling
E-Cadherin	3195s	1:1000		Cell Signalling
N-Cadherin	14215s	1:1000		Cell Signalling
GATA6	ab22600	1:500	1:100	Abcam
HIF1-α	ab51608	1:1000		Abcam
PDX1	5679T	1:500		Thermo Fisher
SMAD4	38454	1:1000		Cell Signalling
TET2	ab124297	1:500		Abcam
Human Mitochondrial Membrane	MAB1273		1:100	Merck
ki67	AB9260		1:1000	Merck
β-catenin	sc-376959		1:100	Santa Cruz
Anti-Mouse HRP	315-035-048	1:5000		Jackson Immuno Research
Anti-Rabbit HRP	7074S	1:5000		Cell Signalling

Antibodies used for Western Blot (WB) and Immunohistochemisty (IHC).

2.3 Generation of Genomic Data from a Cohort of PDAC Patients

A cohort of 141 treatment naive resectable PDAC patients recruited from the Oxford area was used to generate a database of genomic data (referred to as the Oxford Cohort).

2.3.1 Ethics and Patient Material

Formalin fixed paraffin embedded (FFPE) tissue blocks were obtained from the Department of Pathology archive together with clinical follow-up data and diagnostic images at the Oxford University Hospital NHS Trust. Patients had previously provided an informed consent. The present study was approved by the Oxford Radcliffe Biobank, University of Oxford (Project: OCHRe 14/A176). 4um FFPE sections on glass slides were used for histochemical analysis and FFPE scrolls were used for molecular analysis.

The type of pancreatectomy performed was conducted according to international guidelines. Patients included in the present retrospective study had to meet the following criteria: histologically-confirmed PDAC, complete macroscopic surgical resection (R0 or R1), lack of metastatic spread and/or ascites, absence of previous history of malignancy and archived FFPE surgical samples at the Department of Pathology. In total, n = 141 patients were included in the total cohort. The majority of patients were treated with adjuvant chemotherapy consisted of gemcitabine (GEM) as monotherapy at the Oxford University Hospital NHS Trust, Oxford, UK. GEM alone was administered intravenously (dose 1,000 mg/m2 over 30 minutes), at days 1, 8 and 15 (1 cycle) for up to 6 cycles. A few patients received a combination of gemcitabine with capecitabine (GEM-CAP). In this case, GEM was administered as described above and CAP was administered orally (dose 830 mg/m2 twice daily) for

3 weeks followed by 1-week pause.

2.3.2 mRNA Isolation and Microarrays

Isolation of mRNA from FFPE samples was performed by Aline Hubert (Eric O'Neill lab) using Qiagen RNeasy FFPE kits (Table 2.5) according to the standard protocol. Two 4um scrolls were used for each sample, and excess paraffin was removed by xylene washes. RNA was then eluted in 30ul RNase free water. RNA quality and quantity was determined by Agilent Bioanalyzer and Qubit Fluorometer (Thermo Fisher) respectively. Purified RNA was then run on Almac Xcel Arrays. RNA QC was performed by Andrew Blake.

2.3.3 DNA Isolation

DNA was extracted using Qiagen QIAamp DNA FFPE tissue kits (Table 2.5) with Qiagen Deparaffination solution. 500ul Deparaffination solution was added to FFPE tissue scrolls and samples were incubated at 56 °C for 10 mins before cooling to room temperature. 180ul ATL buffer was added and samples were centrifuged for 1 min at 11,000g. 30ul proteinase K was added to the lower, clear phase of the sample before incubating overnight at 56 °C, followed by 1 hour at 90 °C. The remaining protocol was carried out using the standard procedure. DNA was eluted in 50ul Nuclease free water and DNA quantity and quality was determined by Agilent Bioanalyzer and Qubit Fluorometer (Thermo Fisher) respectively. Purified DNA was used for Ion Torrent Sequencing, Methylation and Hydroxymethylation analysis.

2.3.4 Genome Wide Methylation and Hydroxymethylation

In order to determine genome wide hydroxymethylation the CEGX TrueMethyl Array Kit (Table 2.5) was adapted for use with FFPE material. 1ug of DNA in 50ul

ultrapure water was used as starting material. The use of high quality DNA without ethanol or paraffin contaminants was found to be critical for successful completion of the protocol. Additionally, all steps prior to sample oxidation were performed in a room where ethanol had not been sprayed for at least 12 hours to further reduce the risk of ethanol contamination. Low retention pipette tips were used throughout to prevent loss of magnetic beads. DNA was first purified by adding 100ul of Magnetic Bead Binding Solution 1 (MBBS1) and incubating for 20 minutes at room temperature. Tubes were then places on a magnetic rack for 5 minutes to immobilize magnetic beads and the supernatant removed. Beads were washed and aspirated 3 times in 200ul 80% acetonitrile in a fume hood. After the final wash excess acetonitrile was aspirated and pellets were left to air dry for 5 minutes. DNA was denatured by adding 20ul of denaturing solution directly to the bead pellet. Samples were vortexed to ensure proper resuspension of beads and then incubated for 5 minutes at 37 °C before separating on a magnetic rack again. The supernatant containing denatured DNA was then split into two separate 0.2ml PCR tubes. One tube was labelled as the oxBS sample, and the other the BS sample. The oxidative step was carried out on the oxBS sample only.

The oxBS sample was oxidized by adding 1ul of Oxidant solution and incubating at 40 °C for 10 minutes. The sample was then centrifuged at 14,000g for 10 minutes to pellet any precipitate that formed during oxidation and the supernatant transferred to a new tube. Following oxidation and centrifugation the sample should remain an orange colour, indicating excess oxidative reagent in the sample. If the sample is a black or dark brown colour it indicates that the oxidative reagent has been quenched, probably by contaminating alcohols, and the reaction is likely to be incomplete. If this is the case then the sample is discarded.

After the oxidation step, all steps were performed on both the BS and oxBS samples. Samples underwent bisulfite conversion by addition of 30ul bisulfite reagent

and were then placed in a thermocycler. DNA was denatured at 95 °C for 5 minutes, incubated at 60°C for 20 minutes, denatured at 95 °C for 5 minutes, incubated at 60 °C for 40 minutes, denatured at 95 °C for 5 minutes and then incubated at 60 °C for 45 minutes before holding at 20 °C. Tubes were centrifuged at 14,000g for 10 minutes and 40ul of supernatant was transferred to a new 1.5ml eppendorf. DNA was purified by adding 160ul Magnetic Bead Binding Solution 2 (MBBS2) and incubating for 5 minutes. Samples were then placed on magnetic racks for 15 minutes to allow beads to separate. The supernatant was discarded and the pellets washed in 200ul of 70% ethanol. The bead pellet was resuspended in 200ul Desulfonation buffer for 5 mins before placing on the magnetic rack again, aspirating the supernatant, and washing with 200ul 70% ethanol 3 times. After washing and aspirating the bead pellet was allowed to air dry for 15 minutes. DNA was the eluted in 12ul elution buffer. Beads were incubated for 10 mins at room temperature before being separated on a magnetic rack and the supernatant transferred to a new 1.5ml tube. Samples were then run on illumina EPIC arrays by the Oxford Genomics Group, Oxford, UK.

2.3.5 Ion Torrent Sequencing

Ion Torrent Sequencing was performed by Adam Burns (adam.burns@oncology.ox.ac.uk). Genomic mutations in PDAC tumours were determined for a panel of 52 genes commonly mutated in multiple cancers. This focused approach allowed for a higher sequencing depth than if whole exome sequencing had been used. Sequencing libraries were generated using both the Oncomine BRCA Research Assay Chef-Ready kit (Life Technologies, Carlsbad, CA, USA), and the Ion Ampliseq Cancer Hotspot Panel V2 Chef-Ready kit (Life Technologies, Carlsbad, CA, USA). Libraries were barcoded on an Ion Chef instrument (Life Technologies) and sequenced on an Ion Torrent PGM System using Ion 318 sequencing chips (Life Technologies). Data analysis was performed using Torrent Suite Software V.5.2. Only sequences with a p-value <0.001,

sequencing depth >100x and a functional consequence (missense, nonsense or indel) were used. Average sequencing depth after filtering was 474. Mutation rates in KRAS (78%), TP53 (50%), SMAD4 (26%) and CDKN2A (22%) were at similar rates to previous sequencing studies on PDAC tumours.

2.4 Statistical Analysis and Bioinformatics

Unless stated otherwise, all statistical analysis was carried out using R version 3.4.3. In all figures: * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$ and ns = not significant.

2.4.1 Statistical Tests

Students T-Tests were performed in Microsoft Excel for comparison of two groups. When multiple comparisons were made at once the Benjamini-Hochberg (BH) procedure was used with a false discovery rate of 5% to adjust for multiple comparisons. Chi-Square independence tests were used to compare categorical variables between groups, unless an expected value was less than 5, in which case Fisher's exact tests were used. Correlation between two continuous variables was analyzed using Pearson correlation coefficient. Kaplan-Meier curves and Cox Proportional Hazard Regressions were performed using the R Survival package. For IHC quantification of orthotopic tumours Mann-Whitney U Tests were performed in Graphpad Prism 7.0. Graphs were drawn using the ggplots2 package or Graphpad Prism 7.0, and clustered heatmaps drawn using the pheatmap package.

2.4.2 Clinical Datasets

RNA-seq, Methylation 450K and Clinical data was downloaded for Pancreatic Adenocarcinoma (PAAD) by The Cancer Genome Atlas (TCGA) version 2016_1_28 from the Firebrowse portal (Firebrowse.org). A SMAD4 Chip-seq performed in A2780

cells was downloaded from the Gene Expression Omnibus (GEO) accession number GSE27526. Expression data for SMAD4 Tet-off and Tet-on mouse pancreatic cancer cell lines by David et al. (2016) was downloaded from GSE72069.

2.4.3 Methylation/Hydroxymethylation Analysis

EPIC array data was processed using the R package ChAMP (Tian, 2017). Probes with a low detection p-value were removed and samples with 5% of probes with low detection were excluded from analysis. Probes on the XY chromosomes, non-CG probes, those identified as containing SNPs and probes that align to multiple locations were filtered, leaving 625,994 probes for analysis. Differential methylated probe (DMP) analysis was carried out using the champ.DMP() function using an adjusted P value cutoff of 0.05. Due to the presence of negative β-values in the hydroxymethylation dataset, differential hydroxymethylated probe (DhMP) analysis was performed manually by two-tailed t-test followed by BH adjustment for multiple comparisons. Gene set enrichment of DMP/DhMPs was performed using the champ.GSEA() function. To determine the enrichment of probes at different genomic features the CpG.GUI function was used. The distribution of DMP/DhMPs was then compared to the distribution of all probes using chi-squared test. To generate genome tracks for visualization of methylation/hydroxymethylation across genes, the average β-value at for each probe was taken for healthy and tumour samples or different molecular subtypes. These were then uploaded and visualized on the UCSC Genome Browser (http://genome/ucsc.edu/) assemby GRCH37/hg19.

2.4.4 Gene Set Enrichment and Molecular Subtyping

Gene set enrichment of mRNA datasets was performed using the R package GSVA, which can determine the relative enrichment of a gene set across a dataset (Hanzelmann, 2013). Gene sets were downloaded from the Broad Institute MSigDB database

v6.2 (Hallmarks, c2 Curated, c5 Gene ontology and c6 oncogenic signatures). Gene sets significantly associated with a group were determined by two-tailed T-Test followed by BH adjustment. PDAC molecular subtypes were also determined using GSVA. Gene sets were manually created for each subtype using the geneset() and genesetcollection() functions from the GSEABase package. The Basal and Classical signatures from Moffitt et al. (2015) were taken from figure 3A and Gene sets for Collisson subtypes were created using the genes listed in PDAssigner in Figure 1. Squamous, progenitor, ADEX and immunogenic gene sets were created using the core gene programmes listed in the supplementary figures by Bailey et al. (2016) (GP1, GP2, GP6 and GP9). The relative enrichment of each subtypes gene set was then calculated. For each set of subtypes the GSVA results were hierarchically clustered using the pheatmaps package and subtypes were assigned according to which subtype geneset was enriched in each cluster. For determining molecular subtypes in cell lines, normalized ZScores from RNAseq of pancreatic cell lines was downloaded from the Cancer Cell Line Encyclopedia (Barretina et al., 2012). As Exocrine-like and ADEX subtypes have not been reported in cell lines and immunogenic subtypes are dependent on immune cells within the tumour microenvironment these subtypes were excluded from the cell line analysis. The remaining subtype enrichment scores were hierarchically clustered and cell lines assigned a squamous or classical-pancreatic subtype depending on which gene sets were enriched in each cluster.

2.4.5 iCluster

Integrated analysis of mutational, methylation and hydroxymethylation data was performed using the iClusterPlus package (Shen et al., 2009). For gene mutations, binary data was used (1 = mutated, 0 not-mutated) for all genes mutated in over 5% of samples. For methylation and hydroxymethylation data the top 5,000 most variable probes were used. Bayesian information criteria (BIC) modelling using the

tune.iCluster() function was used to determine the optimal number of clusters to explain the variation within the data. The probes and mutations most closely associated with each cluster were shown using the plotHeatmap() function. Significant associations between clusters and mutations or clinical characteristics were determined by Chi-Squared or Fisher's tests.

2.5 Molecular Biology and *in vitro* experiments

2.5.1 Cell Culture

All cell lines (Table 2.1) were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) Fetal Bovine Serum (FBS), 2 mM Glutamine and 100 U/mL Penicillin/Streptomycin (Pen/Strep). Cells were maintained at 37 °C, with 5% CO^2 in a humidified incubator. When confluent cells were split using 0.25% Trypsin-EDTA solution (Life Technologies). Cells were routinely frozen down in freezing medium containing 90% (v/v) FBS, 10% (v/v) Dimethyl sulfoxide (DMSO). Trypsinised cells were pelleted and resuspended in freezing media and aliquoted into 2ml cryovials (Thermo Scientific). Cryovials were placed into a room temperature Nalgene Mr Frosty containing 200mL isopropanol and placed at -80 °C for at least 24 hours prior to long term storage in liquid nitrogen.

2.5.2 Transfection

All cells undergoing transfection were plated in complete DMEM, without antibiotics. All transfections, either with plasmid DNA or siRNA, were done using the Lipofectamine 2000 transfection reagent (Life Technologies) and following the manufacturer's guidelines. This was done at a ratio of 1 ug siRNA to 1 uL of Lipofectamine or 1 ug plasmid DNA to 1 uL of Lipofectamine. All cells were transfected with 100nM siRNA. Appropriate volumes of siRNA or plasmid DNA and Lipofectamine were diluted in

separate tubes in OptiMEM (Life Technologies) and incubated at room temperature for 5 min. Then the contents of the two tubes were mixed together carefully and left at room temperature for 20 min. After incubation the mixture was gently added to the cells. After approximately 6 hours the medium containing the mixture of siRNA or plasmid DNA and Lipofectamine was aspirated and fresh complete DMEM was added in which, unless stated otherwise, the cells were kept until harvesting. For siRNA experiments siLuciferase was used as a control (Eurofins, A4403, GCCAUUCUAUCCUCUAGAGGAUG). Human SMAD4 siRNA was purchased from Dharmacon (M-003902-01-0005).

2.5.3 *in vitro* Drug Treatments

For in vitro drug treatments of cell lines, Panc1, MiaPaca2 and Capan1 cells were treated for 96 hours, and PSN1 and DEC-hTERT cells were treated for 48 hours. Complete DMEM supplemented with L-Ascorbic Acid (Sigma, A4403) and/or Metformin (Abcam, ab146725) were prepared fresh in dH2O or PBS respectively for each day of treatment. Compound C (Santa Cruz, sc-200689) was dissolved in DMSO at a stock solution of 20mM and stored at -20 °C. All drugs used were dissolved in the appropriate buffer and filtered through a 0.4um filter syringe unit to sterilize prior to use. Following treatment cells were harvested and protein, DNA and RNA were isolated for western blot, dot blot and qPCR analysis.

2.5.4 Protein Isolation, SDS-Page and Western Blotting

Harvested cell pellets were lysed in Laemmli buffer (Table 2.4) and placed on a rotor at 4 °C for 15 minutes. Lysates were homogenised using a 23 gauge syringe needle and centrifuged for 10 mins at 13,000 rpm. Protein concentration was determined using a NanoDrop 1000 (Thermo Fisher) by measuring the absorbance of the samples at a wavelength of 260 nm, with an optical density of 1 equivalent 50 ug/mL. 1 volume of

2X SDS-PAGE sample buffer (Table 2.4) was added to each sample, which were then boiled for 15 min at 100 °C before loading onto a SDS-PAGE gel for SDS-PAGE gel electrophoresis. Lysates were loaded into NuPAGE pre-cast 4-12% Bis-Tris protein gels (Life Technologies) in Novex mini gel tanks filled with 1X NuPAGE MOPS SDS running buffer (Table 2.4). SeeBlue Plus 2 Pre-stained protein standard (Life Technologies) was used as a protein marker. Gels were run at a constant voltage of 120V for 2 hours using PowerPac HC High-current power supply (Bio-Rad). Following gel electrophoresis, protein samples were transferred from the gel onto an Immobilon-P PVDF membrane (Millipore). The membrane was activated using methanol before the transfer was performed using the Bio-Rad mini gel blotting system, in which the gel and membrane were sandwiched between two sheets of Whatman paper and compacted using sponges. The transfer was performed in transfer buffer (Table 2.4) at a voltage of 100 V for 80 min using PowerPacTM HC High-current power supply (Bio-Rad). Following transfer, PVDF membranes were blocked for 1 hour on a SSL4 See saw rocker (Stuart) at room temperature in 5% skimmed milk (w/v) in PBS-Tween (PBS-T, Table 2.4). After blocking, the primary antibody was added to the membrane at the appropriate concentration (Table 2.6) in 5% skimmed milk PBS-T. All primary antibodies were incubated with the membrane overnight at 4 °C.

Following primary antibody incubation, membranes were washed (3 x 5 min) in PBS-T on a See-saw rocker before being incubated with secondary HRP conjugated antibody (Table 2.6). Membranes were incubated at room temperature for 45 min prior to being washed (3 x 5 min) in PBS-T. The PVDF membranes were then exposed to either Enhanced chemiluminescence (ECL) solution (Thermo Scientific) for normal exposure or ECL advance (GE Healthcare) for low abundance proteins, according to manufacturer's protocol. Membranes were exposed to X-ray film (Scientific Laboratories Supplies) and developed using a XoGraph processor (XoGraph Imaging Systems) before being scanned onto a computer. In order to allow membranes to be

probed with multiple primary antibodies, membranes were de-activated/re-blocked in 5% milk supplemented with 1:500 sodium azide 10% (w/v) for 30 min on a See saw rocker. Further to that, new primary antibody incubation was carried out.

2.5.5 DNA Isolation and Dot Blotting

Cell pellets were lysed in DNA lysis buffer (Table 2.4) and incubated at 55 °C whilst shaking for 72 hours. Samples were then centrifuged for 10 mins at 13,000 rpm and the supernatant transferred to a new Eppendorf in order to remove any residual protein or cell aggregates. Genomic DNA (gDNA) was precipitated by adding 1 volume of isopropanol and incubating at room temperature for 5 minutes before centrifuging at 12,000g for 10 minutes at 4 °C. The supernatant was aspirated and DNA pellets were washed by resuspending in 70% ethanol and centrifuging at 12,000g for 5 minutes at 4 °C. After aspirating the supernatant again excess ethanol was allowed to evaporate and the pellets were resuspended in 50-100ul nuclease free water. RNAse A digestion was carried out at 37 °C for 2 hours. gDNA was sonicated in a Diagenode biorupter plus for 20 cycles of 30 seconds on, 30 seconds off at high power. DNA was then quantified using a NanoDrop 1000 (Thermo Fisher). For dot blots, DNA was normalized and then denatured by adding 1 volume DNA denaturing buffer and heating at 95 °C for 10 minutes. 2 volumes 20X saline-sodium citrate (SSC) buffer was added and samples were chilled on ice for 5 minutes. Serial dilutions of DNA were directly pipetted onto Amersham Hybond-N+ blotting paper (GE Healthcare) and UV crosslinked at 1200J/m2. Membranes were blocked with 5% milk/PBST and incubated with 1:1000 5'hmC primary antibody (Table 2.6) overnight at 4 °C. Membranes were then washed and developed in the same method used for western blot as above. To visualize total DNA membranes were stained with 0.1% Methylene blue. The 5'hmC signal was quantified using imageJ and normalized to total DNA level.

2.5.6 RNA Isolation and qPCR

RNA isolation and cDNA synthesis was performed using Power SYBR Green Cells-to-Ct Kit (Table 2.5) following the standard protocol. 50,000 cells were lysed in 25uL Lysis solution containing DNase, and 5ul of this lysis used for cDNA synthesis. Target gene expression was assessed using Power SYBR Green PCR Master Mix (Thermo) using primers designed by PrimerBank (Zaidi et al., 2010b). All reactions were performed in duplicate and ran for 45 cycles. Fold change in target mRNA expression was calculated based on the cycle threshold (CT) method after normalization to GAPDH expression.

2.6 *in vivo* Experiments and Histological Analysis

2.6.1 Orthotopic Injection of PDAC Cell Lines into Mice

All animal studies were performed in accordance with the UK Animals (Scientific Procedures) Act of 1986 under licences approved by the UK Home Office and with the approval of the University of Oxford ethical review committee. Animals were housed in environmentally enriched (activity wheel or crawl ball, nesting material and chew sticks) IVC cages in groups of 5 per cage, in a 12 hour day-night cycle facility maintained at 22°C in 50% humidity. All mice had *ad libitum* access to certified food and tap water.

In order to establish pancreatic orthotopic tumours, Human pancreatic cell lines (PANC1, MiaPaca2 and PSN1) were resuspended in 50-50 matrigel-DMEM solution and kept at 4C until the injection into the pancreas. At the time of the injection, 20ul with a total of 500,000 cells were injected in the tail portion of the pancreas.

For orthotopic surgery athymic nude-Fox1nu mice were used. An incision was made in the left abdominal side, the tail region of the pancreas was exteriorized, and

the appropriate cell suspension was injected via 26G syringe. Successful injection was veried by the appearance of a wheal at the injection site with no leakage through the pancreatic capsule. The abdominal wall was sutured with absorbable 4.0 Vicryl suture (Ethicon), and the skin was closed with wound clips. Tumour growth were monitored post surgery using 4.7T MRI scan (Bruker BioSpin GmbH, Germany). In the treatment cohorts, Metformin (Abcam) was provided in the drinking water (2.5mg/ml) from day 1 (11 days after the surgery). Sodium Ascorbate (Sigma-Aldrich) was administrated (2g/kg in PBS) via intraperitoneal injection daily from day 1 until day 5 concurrently with Metformin. Mice were euthanized at day 5 of the treatment.

2.6.2 Tissue Processing and Embedding

Mouse tissues were dissected and fixed in 10% neutral buffered formalin (NBF) overnight before storing in 70% ethanol for 24 hours. Tissues were processed using a STP20 tissue processor (Thermo) and embedded in paraffin wax using HistoStar tissue embedder (Thermo). 4um sections were cut on positively charged glass slides for histological analysis.

2.6.3 Histological Analysis

FFPE Sections on glass slides were de-waxed in two washes of Citroclear and rehydrated in graded alcohol concentrations for 3 mins each (100%, 100%, 90%, 70%, and 50%), followed by two washes in PBS. Slides were then stained by Immunohistochemistry (IHC), Haematoxylin and Eosin (H& E) or Masson's Trichrome. Following staining, slides were dehydrated in ethanol gradients (50%, 70%, 90%, 100%, 100%) for 30 seconds each and cleared in two xylene washes for 1 minute each. Sections were mounted in Di-n-Butyl phthalate in Xylene (DPX) and scanned using an Aperio ScanScope slide scanner.

2.6.3.1 Immunohistochemisty (IHC)

Antibody dilution, antigen retrieval temperature and retrieval buffer was optimized for each antibody. In all cases retrieval conditions of 125 °C for 2 mins in Tris-EDTA buffer (pH 9) was found to give the lowest level of background staining. Antigen retrieval was performed using a pressure cooker. Slides were headed up to 125 °C for 2 minutes, then allowed to cool to 60 °C before placing on ice for an hour. Sections were then washed in PBS for 3X 5 mins and blocked using Dako envision+ Dual Endogenous Enzyme Block for 1 hour. Primary antibodies (table 2.6) were incubated overnight at 4 °C. Slides were then washed in PBS three times and then incubated in secondary rabbit-HRP (Dako) for 30 minutes at room temperature followed by 3 PBS washes. Slides were developed with 3,3'-diaminobenzidine (DAB) solution (Dako) and inactivated in water before counter staining in haematoxylin for 30 seconds. Slides were then rinsed, dehydrated, cleared and mounted.

Human tumours stained with 5'hmC were quantified by a pathologist (Robert Goldin; r.goldin@imperial.ac.uk). For orthotopic pancreatic cell line tumours 3-5 randomly chosen 300um squares were quantified using imageJ. Only tumour cells were quantified (cells with large irregular nuclei). For ki67 staining the proportion of cells with positive nuclear staining was quantified. For GATA6, β-catenin and 5'hmC the nuclear staining intensity was measured for all tumour cells within each 300um square. Thresholds for weak, moderate and strong intensity of staining were set and the proportion of each within the square determined using the nuclear intensity measurements. Mann-whitney U-tests were performed on the amalgamated measurements for each cell line and treatment. Presented data shows average proportions for each group with error bars showing standard deviation. Error bars are shown on the negative axis only to make data less cluttered.

2.6.3.2 Haematoxylin and Eosin (H&E)

Rehydrated slides were placed in freshly filtered Harris Haemotoxylin for 2 minutes, then washed in running tap water for 5 minutes. Sections were differentiated with 2 dips in 1% Acid Alcohol solution followed by 10 dips in dH20. Sections were then blued by 10 dips in Scott's Tap water, 10 dips in dH20 and 10 dips in 0.5% lithium carbonate. Slides were washed under running tap water and then dehydrated in ethanol gradients. Dehydrated slides were then placed in eosin solution for 1 minute. Excess eosin was removed by 4 washes in 90-100% ethanol for 30 seconds each and the dehydrated slides were then cleared and mounted.

2.6.3.3 Masson's Trichrome

Collagen was stained using Masson's Trichrome Stain Kit (Table 2.5). Slides were immersed in preheated Bouin's solution at 56 °C for 15 minutes to mordant. Sections were then washed in running tap water for 5 minutes before staining with Weigert's Iron Haemotoxylin solution for 5 minutes. Sections were washed in running tap water again then stained with Biebrich Scarlet-Acid Fuchsin for 5 minutes. Slides were rinsed with dH2O and placed in Phosphotungstic/Phosphomolybdic acid solution for 5 minutes. Collagen was then stained with Aniline Blue solution for 5 minutes, followed by washing in 1% Acetic Acid for 2 minutes. Slides were then rinsed in dH20, dehydrated, cleared and mounted in DPX. The proportion of tumours with positive collagen staining was determined using imageJ. Three 300um squares were used for each tumour. The blue channel was deconvoluted and a threshold was set for positive staining in order to determine the proportion of positive staining.

2.6.4 RNA-seq on Orthotopic Tumours

DNA, RNA and miRNAs were simultaneously purified from orthotopic tumours using Qiagen AllPrep DNA/RNA/miRNA Universal Kits (Table 2.5). 600ul RLT lysis

buffer was added to ∼30mg of flash frozen tissue which was then disrupted using sterile TaKaRa biomashers. Lysates were then transferred to Qiagen QIAshredders and centrifuged at full speed for 2 minutes to homogenize. The remaining steps were performed according to the standard protocol. DNA was eluted in 50ul elution buffer, and RNA was eluted in 50ul RNAse free water. Quantity and quality of DNA and RNA was determined using Nanodrop, and RNA integrity was analyzed using Agilent Bioanalyzer.

Paired 3'RNA-seq was performed on isolated RNA by the Oxford Genomic Centre, Oxford, UK. Following sequencing, raw fastq reads were aligned to the human genome (hg19) using RNA STAR on the Galaxy web platform (Afgan et al., 2016). Aligned BAM files were then further processed using R. A gene model was generated from the UCSC hg19 known gene database (R package TxDb.Hsapients.UCSC.hg19.knownGene). Genes were then grouped by transcript using the transcriptsBy(, "gene") function in the GenomicFeatures package (Lawrence et al., 2013). A count matrix was then generated by counting the number of reads within each gene region using the summarizeOverlaps() function in the GenomicAlignments package (Lawrence et al., 2013). This count matrix was converted to a DESeqDataSet object using the DESeq2 package and read counts were normalized according to size factors determined by DESeq2 (Anders and Huber, 2010). As 3'RNA-seq only gives one read per transcript there is no need to normalize counts according to transcript size. Therefore, the log2 normalized counts were used for further analysis. Entrez gene IDs were converted to gene symbols using the MapIds() function from the AnnotationDbi package in combination with the Homo.sapiens package. To deconvolute human and mouse specific reads, raw Fastq files were also aligned to the mouse genome (mm10). Reads that aligned to both genomes were then attributed to the genome which they most closely aligned to using the XenoFilteR package (Kluin et al., 2018). Filtered BAM files were then processed in the same way as described above, except that for the mouse reads a mouse gene

model was generated using the TxDb.Mmusculus.UCSC.mm10.knownGene package.

Chapter 3

Whole Genome 5'hmC Profiling in a Cohort of PDAC Patients

3.1 Introduction

Loss of proper differentiation enables cancer cells with stem-like properties to accumulate and drives tumour metastasis and recurrence (Ayob and Ramasamy, 2018). In pancreatic cancer, multiple genomic analysis using mRNA datasets have identified highly aggressive subtypes of tumours characterized by loss of endodermal differentiation markers such as GATA6 and PDX1. (Collisson et al., 2019), suggesting that loss of tumour differentiation is a driving force in PDAC progression. Despite the inherent heterogeneity in the molecular basis of PDAC, there is little genetic heterogeneity between different PDAC tumours, even when comparing metastasis to primary tumours (Reiter et al., 2018). Conversely, there is widespread epigenetic reprogramming association with progression to more aggressive phenotypes (McDonald et al., 2017), suggesting that tumour metastasis, de-differentiation and more aggressive molecular subtypes are likely to be epigenetically driven.

TET-mediated DNA hydroxymethylation (5'hmC) functions as an intermediate

in the active DNA demethylation process, but is increasingly appreciated as an epigenetic mark in its own right by antagonizing DNA methylation (5'mC). 5'hmC is dynamically regulated during pancreatic cell differentiation, and accumulates at developmental genes such as PDX1 during their activation, however it is unknown what role 5'hmC plays during pancreatic tumorigenesis.

3.1.1 Bisulfite Conversion of DNA for Detection of Methylation

Bisulfite sequencing is the current "gold standard" for the detection of DNA methylation and has been used since 1992 to give a qualitative and quantitative overview of 5'mC at base pair resolution (Frommer et al., 1992). Treatment of DNA with sodium bisulfite deaminates unmodified cytosine to uracil, but leaves methylated cytosines intact. In a subsequent PCR reaction, uracils are amplified as thymines, while modified cytosines that were protected from deamination remain as cytosine (Fig. 3.1). By comparing DNA sequences before and after bisulfite conversion the methylation status of cytosines can then be determined. However, 5'hmC can also protect cytosines from deamination during bisulfite conversion (Huang et al., 2010). Therefore, conventional bisulfite sequencing cannot distinguish between these two antagonistic epignetic marks and all methylation studies using this technique are actually a combination of 5'mC and 5'hmC.

3.1.2 The oxBS method for detection of 5'hmC

Potassium Perruthenate can selectively oxidise 5'hmC to 5-formylcytosine (5'fC), while leaving 5'mC unchanged. The oxBS method utilizes this reaction in combination with bisulfite sequencing in order to distinguish 5'mC and 5'hmC, as 5'fC is sensitive to deamination by sodium bisulfite (Fig. 3.1). Samples are split into two,

with one undergoing conventional bisulfite conversion (BS) and the other undergoing an oxidative step prior to bisulfite treatment (oxBS). "Methylated" cytosines in the conventional BS sample are a combination of 5'mC and 5'hmC, while in the oxBS sample 5'hmC has been oxidized and only 5'mC remains. 5'hmC can then be quantified at base pair resolution through a simple subtraction of BS minus oxBS (Booth et al., 2012). This method has been used to successfully investigate 5'hmC in cell lines and a small number of studies using fresh frozen tissue (Johnson et al., 2016), but has yet to be successfully adapted for use with formalin fixed, paraffin embedded (FFPE) tissue, which is the standard method for storage of patient material.

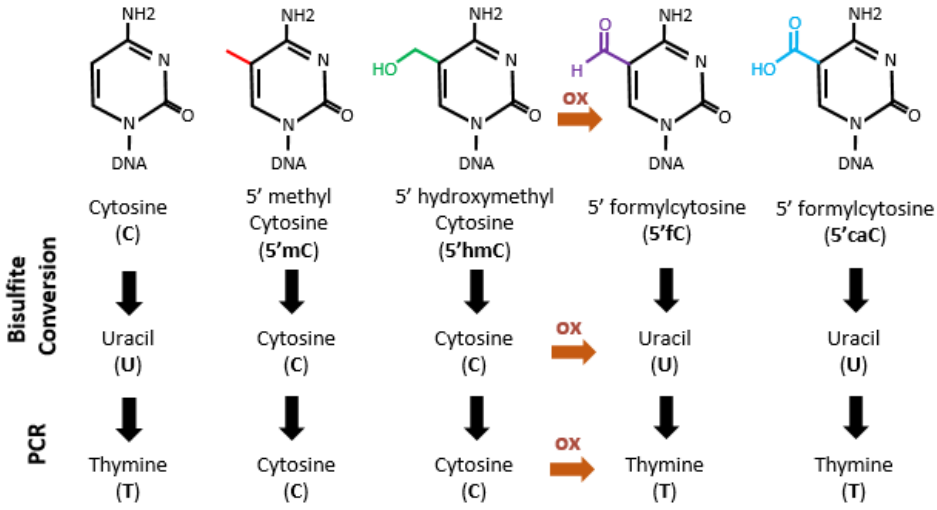


Figure 3.1: **Bisulfite Conversion Effects on Cytosine Modifications**. Diagram showing changes to cytosine residues following bisulfite conversion and subsequent PCR. Unmodified cytosines, 5'fC and 5'caC are deaminated to uracil by sodium bisulfite and amplified as thymines. 5'mC and 5'hmC are protected from deamination and remain as cytosines following PCR. Oxidation of 5'hmC by potassium perruthenate to 5'fC (ox) allows for 5'mC and 5'hmC to be distinguished.

3.1.3 Aims

As 5'hmC correlates with the activation of differentiation genes during pancreatic development, loss of 5'hmC may result, or coincide, with reduced expression of these

genes, resulting in more aggressive tumours. In the following chapter, the correlation between 5'hmC and different clinical characteristics will be investigated using a cohort of FFPE pancreatic tumours. Global 5'hmC changes will be determined using immunohistochemisty (IHC) and the oxBS method will be adapted for use with FFPE material for the first time; providing base pair resolution of 5'hmC abundance genome-wide. Additionally, a new multi-genomic database will be generated using these tumours containing mutational, transcriptomic and epigenetic information in order to better enable research into pancreatic cancer.

3.2 Results

3.2.1 Generation of Genomic Data from a Cohort of FFPE PDAC Patients - The Oxford Cohort

Genomic studies of cancer have greatly increased the understanding of how different tumours progress by outlining the similarities and differences between patient tumours at the molecular level. With this in mind, a new cohort dataset (referred to henceforth as the Oxford Cohort) was created. The creation and analysis of this dataset forms the initial focus of this chapter. Generation of genomic data was highly collaborative: Isolation of DNA for mutational sequencing and methylation/hydroxymethylation arrays was performed by the author, along with adaption of the oxBS method for use with FFPE material (outlined later in this chapter). Mutational sequencing was performed by Adam Burns from the lab of Anna Schuh, and isolation of mRNA for Agilent arrays was performed by Aline Hubert (O'Neill Lab). Quality control (QC) of mRNA and epigenetic arrays was performed by Andrew Blake from the Cancer Centre Bioinformatics Core Facility (Oxford department of oncology). All data analysis presented here was performed by the author. The overlap between different types of genomic data within the cohort is shown in Figure 3.2A.

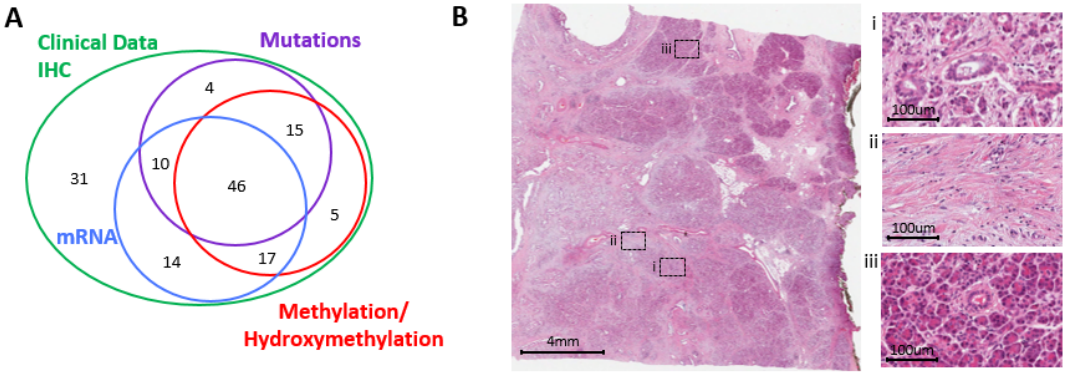


Figure 3.2: **The Oxford Cohort**. **A**) Overlap between different data platforms within the cohort. The cohort consists of 143 patients, but a complete genomic dataset is only available for 46 patients. **B**) Example H&E staining of whole pancreatectomy sections used for analysis. Right, regions of (i) tumour, (ii) stromal and (iii) healthy adjacent tissue within the sections.

The Oxford cohort consists of 143 resectable PDAC Patients recruited from the Oxford area. Patients were treatment naive at the time of surgery and the majority of patients went on to be treated with Gemcitabine, with a small number treated with 5'FU, Folfirinox or Capecitabine. 5 year survival was 17.7%; in line with previous studies on resectable PDAC (Lewis et al., 2013), and 80% of patients recurred within 5 years. Clinical data was retrospectively retrieved by Somnath Mukherjee, Emmanouil Fokas and Ahmad Sabbagh (Oxford University). Multivariate analysis using Cox proportional hazard model was used to determine which clinical factors were important for patient survival. T-Stage, N-Stage, incomplete resection of the tumour (resection margin), venous or lymphatic invasion, tumour recurrence and the site of recurrence, but not gender or tumour grade, were prognostic indicators of survival (Table 3.1).

Table 3.1: **Indicators of Survival in the Oxford Cohort**

Clinicopathological Feature	**%**	**HR**	**95% CI**	**P Value**
Gender				
Female	46	1		
Male	54	0.84	0.53-1.22	0.349
T Stage				
T1	9	1		
T2	13	2.42	0.92-6.38	0.073
T3	72	3.25	1.42-7.46	**0.005**
T4	6	3.41	1.1-10.6	**0.034**
Grade				
G1	6	1		
G2	56	1.04	0.47-2.29	0.912
G3	26	1.83	0.82-4.07	0.138
G4	2	0.51	0.13-2.93	0.537
N Stage				
N0	23	1		
N1+	77	2.63	1.56-4.43	**0.0002**
Resection Margin				
R0	38	1		
R1+	62	1.91	1.27-2.87	**0.002**
Lymphatic Invasion				
No	24	1		
Yes	76	2.26	1.37-3.72	**0.0014**
Vascular Invasion				
No	32	1		
Yes	68	2.10	1.37-3.22	**0.0007**
Recurrence				
No	34	1		
Yes	66	2.84	1.37-3.22	**0.00001**
Site of Recurrence				
Local	21	1		
Distant	29	1.55	0.82-2.93	0.17
Both	50	2.56	1.43-4.59	**0.0016**

Clinicopathological features of the cohort and their association with overall survival. The percentage of patients in each group (%), Hazard ratios (HR) and 95% confidence intervals (95% CI) are shown. Significant P values are highlighted in bold.

Whole pancreatectomy sections were used for analysis, which contain a mix of tumour, stromal and healthy adjacent tissue (Fig. 3.2B). This makes the dataset comparable to other genomic datasets using PDAC tissue, including The Cancer Genome Atlas (TCGA) Pancreatic Adenocarcinoma (PAAD) dataset, as well as the datasets used by Collisson et al. (2011), Moffitt et al. (2015) and Bailey et al. (2016) for initial descriptions of PDAC molecular subtypes. This gives an overview of the whole tissue microenvironment. However, the extensive stromal reaction characteristic of PDAC makes it difficult to dissect out signals specific to tumour cells. For three patients, a tissue sample was also taken far from the tumour site as a healthy control. Haemotoxylin and eosin (H&E) staining of non-affected sections showed histology consistent with healthy pancreatic tissue, with no evidence of inflammation, acinar-ductal metaplasia or PanIN lesions.

Following surgery, resected tissue was formalin fixed and paraffin embedded (FFPE). Typically, fresh frozen tissue is preferable to FFPE for molecular analysis as it provides a cleaner signal with less background noise (Robbe et al., 2018). However, FFPE is the current "gold standard" for the storage of patient tissue. Therefore, the use of FFPE tissue rather than fresh frozen makes it more likely that genomic analysis will be clinically translatable.

3.2.2 Total 5'hmC levels are reduced during PDAC tumorigenesis

Firstly, in order to determine if 5'hmC is altered in pancreatic cancer on a global scale, total nuclear 5'hmC levels were examined within the Oxford Cohort by IHC. Tumour sections on glass slides were available for 125 patients, as well as the 3 healthy controls. Staining of healthy tissue showed moderate to strong positivity throughout the nuclei of the pancreas, with the highest levels of 5'hmC present in the islets, followed by acinar cells and the lowest levels in the ducts. The same pattern was seen

in histologically healthy tissue adjacent to tumours (healthy adjacent) found in 38 whole pancreatectomy sections (Fig. 3.3).

A number of preneoplasic PanIN lesions were identified within sections by a pathologist, 12 of which were low grade and 10 were high grade. Interestingly, 50% of low grade PanINs displayed weak or negative staining, that was increased to 70% in high grade PanINs. Across all stained tumours, only 30% maintained moderate levels of 5'hmC within neoplastic cells. 30% of tumours had weak levels of 5'hmC, while the remaining 40% of tumours were negative for any 5'hmC staining within neoplastic cells (Fig. 3.3). Moderate staining within healthy adjacent tissue or stromal cells was used as an internal control to ensure that this loss of staining was not due to incorrect staining procedure. This indicates that total 5'hmC levels are decreasing during PDAC formation; beginning at the initial stages of PanIN formation and becoming more widespread as tumours progress.

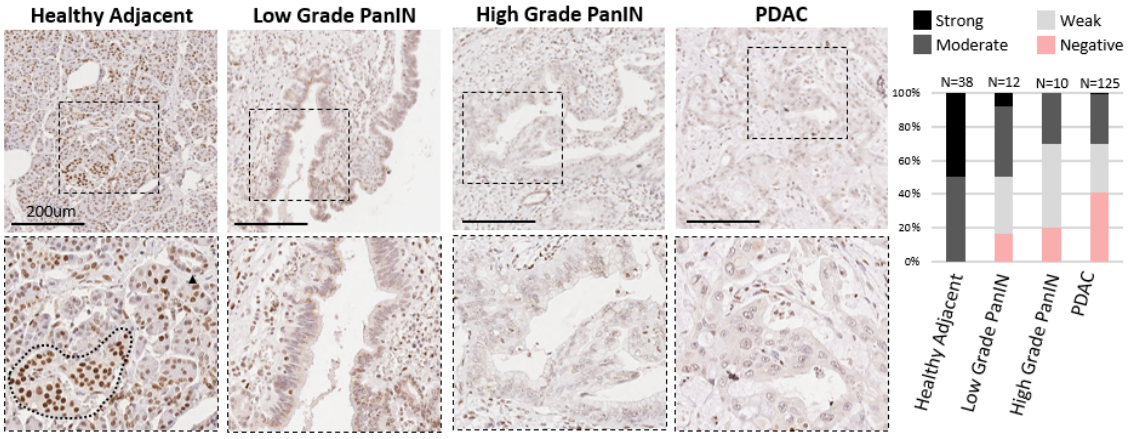


Figure 3.3: **Total 5'hmC Levels Decrease During PDAC Tumorigenesis**. IHC for 5'hmC in pancreatectomy sections. 5'hmC staining intensity was quantified in healthy adjacent tissue, low and high grade PanINs and neoplastic cells (PDAC). 5'hmC staining is progressively lost during PDAC progression.

3.2.3 Detection of genome wide 5'hmC in FFPE Patient Tissue

To determine where in the genome 5'hmC is being lost, the oxBS method was adapted for use with FFPE tissue. The sample workflow is shown in (Figure 3.4A). DNA was isolated from FFPE scrolls and split into two 500ng samples. One sample underwent conventional bisulfite conversion (BS sample) and the other was oxidised prior to bisulfite treatment (oxBS sample). Modified cytosines present in both BS and oxBS samples were protected from conversion by 5'mC and 5'hmC, while modified cytosines present in only the BS sample were protected by 5'mC only. Therefore, the oxBS sample represents the true 5'mC status, while 5'hmC can be determined as the difference between BS and oxBS.

The potassium perruthenate oxidative reagent is highly sensitive and can become easily quenched by contaminants such as ethanol or paraffin, that can result in compromised conversion of DNA. Samples in which excess oxidative reagent is still present after oxidation will maintain an orange colour, while samples that become quenched will turn a grey/black colour (Fig. 3.4B). Given that FFPE samples are stored in paraffin, isolation of high purity DNA free from contaminants was found to be the most important step for oxidative success. Numerous de-paraffinization methods were tested, including Xylene, Qiagen de-parrafination solution and ultrasonic waves (Covaris). Qiagen de-paraffinization solution was more effective at removing excess paraffin from samples and provided a higher yield and purity of DNA. In order to further reduce the risk of contaminants, all steps after DNA isolation until the oxidation reaction were carried out in a clean room where ethanol had not been sprayed for the past 12 hours. During initial attempts, only about 50% of samples successfully passed the oxidative step. After protocol optimization, this was increased to 100%. The complete protocol is available in the methods section of this thesis.

Detection of methylation was performed by running samples on infinium EPIC

arrays. These arrays consist of over 850,000 probes that each measure the methylation status at a single cytosine. The beta (β)-value associated with each probe reflects the extent of methylation of that cytosine, with 0 being completely unmethylated and 1 fully methylated. 5'hmC is quantified as the change in β-value following oxidation of DNA ($\Delta\beta$). Raw density plots of BS and oxBS samples showed a clear drop in methylation signal in samples that underwent oxidation, consistent with a loss of signal due to conversion of 5'hmC (Fig. 3.4C).

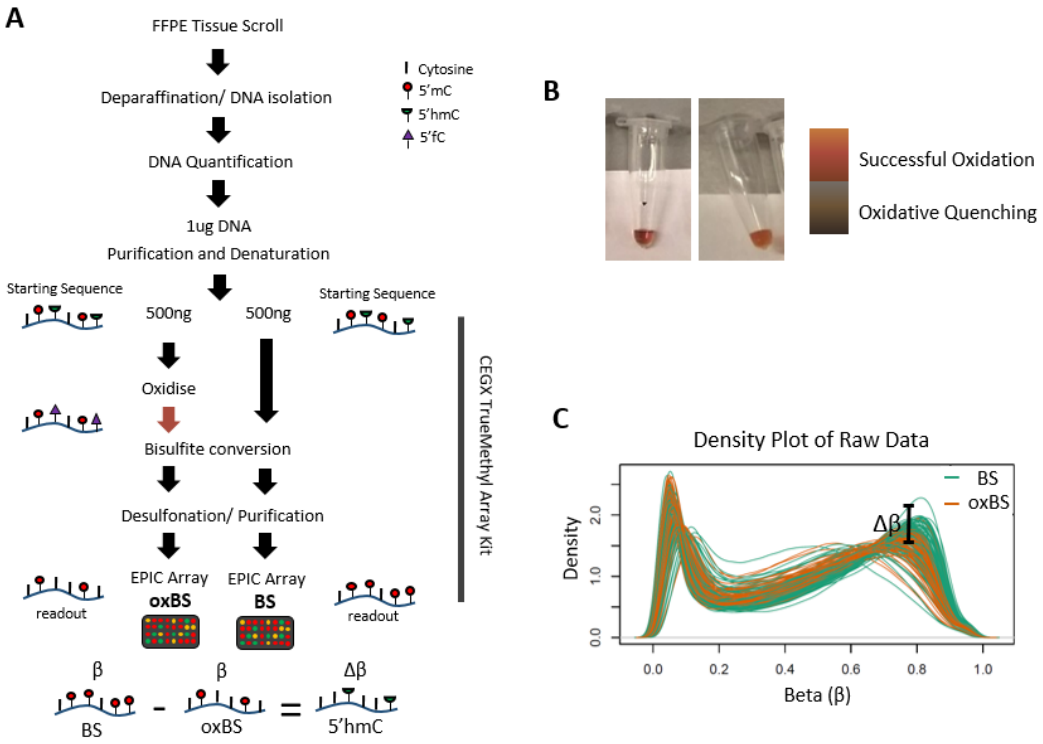


Figure 3.4: **Adaption of oxBS Method to FFPE**. **A**) Work flow for oxBS method. DNA is split into two 500ng samples, one of which undergoes conventional bisulfite treatment (BS), while the other is oxidized prior to bisulfite treatment (oxBS). The full protocol can be found in the methods section. The schematic on either side of the work flow indicates what happens to modified cytosine residues at each stage. **B**) Example of samples following successful oxidation. The orange colour indicates residual Potassium Perruthenate oxidative reagent, meaning the reaction has not been quenched. **C**) Raw density plots from EPIC arrays for BS and oxBS samples. Oxidation of 5'hmC results in a drop in signal for oxBS samples.

To get an overview of the differences between BS and oxBS samples, the average

β-value across all tumour samples was taken for each group. BS and oxBS samples were highly correlated with one another (C=0.991), suggesting that the inability to separate 5'mC and 5'hmC in conventional methylation studies using bisulfite sequencing has a limited effect on results (Fig. 3.5A). 5'hmC levels can be quantified as the differences between BS and oxBS samples ($\Delta\beta$). This method produces both positive and negative $\Delta\beta$-values whereby positive values represent regions of potential 5'hmC and negative values likely reflect background noise (Stewart et al., 2015). The average $\Delta\beta$ for probes across all tumour samples had far more positive values than negative (Fig. 3.5B). Furthermore, almost 50,000 probes have $\Delta\beta$-values over 0.05, while only 40 probes had values below -0.05. This suggests that the positive signal being detected is due to 5'hmC, rather than due to background noise.

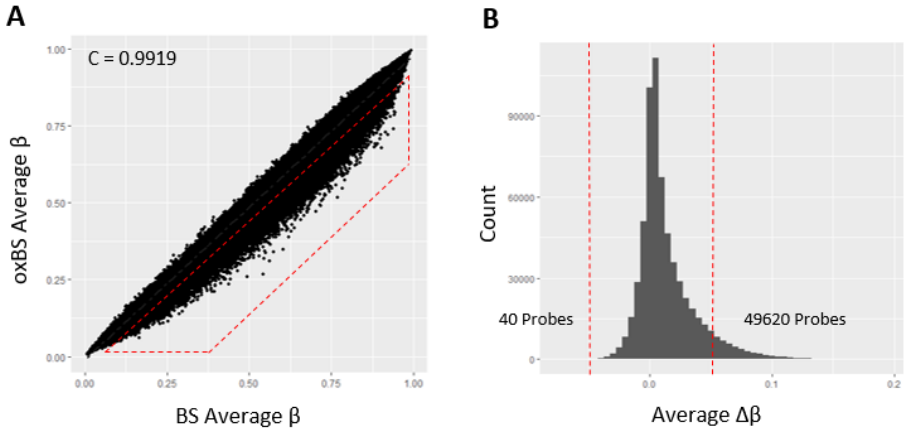


Figure 3.5: **Detection of Methylation and Hydroxymethylation FFPE**. **A**) Correlation between average probe β-values in BS and oxBS samples. Lower signal in oxBS sample relative to BS is due to loss of 5'hmC signal (red box). **B** Average $\Delta\beta$ across all probes. Positive $\Delta\beta$ values indicate hydroxymethylated probes, while negative $\Delta\beta$ values are due to background noise. Only 40 probes have $\Delta\beta$ below -0.05, while almost 50,000 probes had $\Delta\beta$ values above 0.05.

3.2.4 Pancreatic Tumours are Losing 5'hmC at PDAC Specific Loci

In order to determine where in the genome 5'hmC is being lost during PDAC tumorigenesis, differential hydroxymethylation analysis was performed between healthy and tumour samples. In almost all tumour specific differentially hydroxymethylated probes (DMhP) a decrease in 5'hmC was seen (Fig. 3.6A). This is consistent with IHC results that showed that total 5'hmC levels are reduced in PDAC tumours (Fig. 3.3). Gene set enrichment analysis (GSEA) of DMhPs showed that 5'hmC is specifically being lost at genes that are downregulated in PDAC, as well as a number of pathways critical to PDAC progression, including the MAPK pathway and targets of TP53 and TGFβ1. Genes in the insulin pathway, that have been shown to have increased 5'hmC during pancreatic differentiation (Li et al., 2018), were also highly enriched (Fig. 3.6B). Clustering of tumour specific probes revealed that patients typically had low, intermediate or high levels of hydroxymethylation across all probes. Patients that maintained the highest level of hydroxymethylation had a better disease free survival, consistent with the idea that 5'hmC has a tumour suppressive function in PDAC (Fig. 3.6C).

3.2.5 Loss of 5'hmC Correlates With Increased 5'mC

To get a better understanding of what the functional consequences of a reduction in 5'hmC in the PDAC genome might be, the distribution of the most highly hydroxymethylated probes was examined (average $\Delta\beta$ >0.1 across all samples). 5'hmC was depleted at CpG islands, that typically house regulatory elements such as promoters or enhancers and highly enriched at CpG shores ($\pm$ 2kb directly adjacent to the CpG island) (Fig. 3.7A). This pattern has been described previously (Stewart et al., 2015) and is similar to observations by Li et al. (2018) that 5'hmC accumulates adjacent to

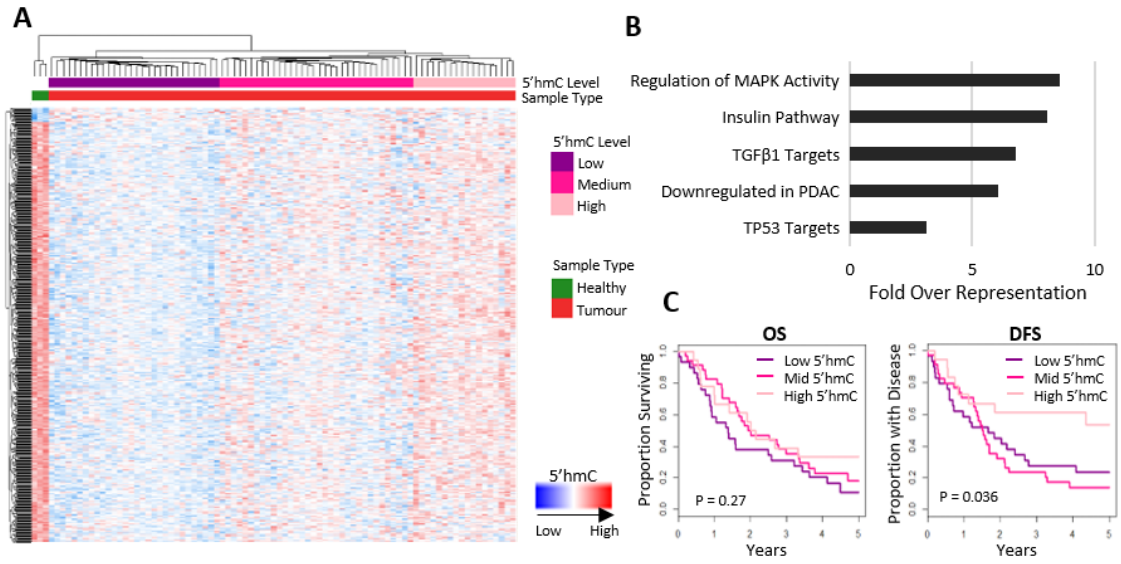


Figure 3.6: **PDAC Tumours Lose 5'hmC at PDAC Specific Loci**. **A**) Clustered heatmap showing deferentially hydroxymethylated probes (DMhPs) between healthy and tumour samples. Patients typically had low, medium or high levels of 5'hmC accross all probes. **B**) Gene set enrichment for tumour specific DMhPs. **C** Kaplan Meier plots for overall and disease free survival for patients with low, medium and high levels of 5'hmC. Patients with the highest levels of 5'hmC had better disease free survival.

regulatory regions in pancreatic cells. These 5'hmC peaks are important for keeping regulatory regions methylation free and their loss can result in epigenetic silencing due to 5'mC spreading into the region (Li et al., 2018)(Wiehle et al., 2016)(Dawlaty et al., 2013). Therefore, the loss of 5'hmC at CpG shores in PDAC tumours may open up CpG islands to become methylated. To determine if this may be the case, the relationship between average 5'mC at CpG islands and 5'hmC at each islands shore was examined in healthy and tumour samples (Fig. 3.7B). This revealed a highly significant negative correlation between the two; supporting the hypothesis that 5'hmC is indeed be protecting CpG islands from becoming methylated. This also further highlighted the drop in 5'hmC associated with PDAC tumours as there were far fewer CpG shores with high levels of 5'hmC in tumour samples relative to healthy. Manual examination of genes known to be methylated in PDAC (Nuclear Receptor Subfamily 5 Group A Member 2 [NR5A2] and Growth Factor Receptor Bound Pro-

tein 10 [GRB10]) showed a depletion of hydroxymethylation peaks adjacent to the promoter region, that correlated with an increase in methylation across the promoter (Fig. 3.7C). These results suggest that the protective 5'hmC peaks that maintain active regulatory regions in differentiated pancreatic cells described Li et al. (2018) are being lost during pancreatic tumorigenesis, opening up these regions to becoming epigenetically silenced.

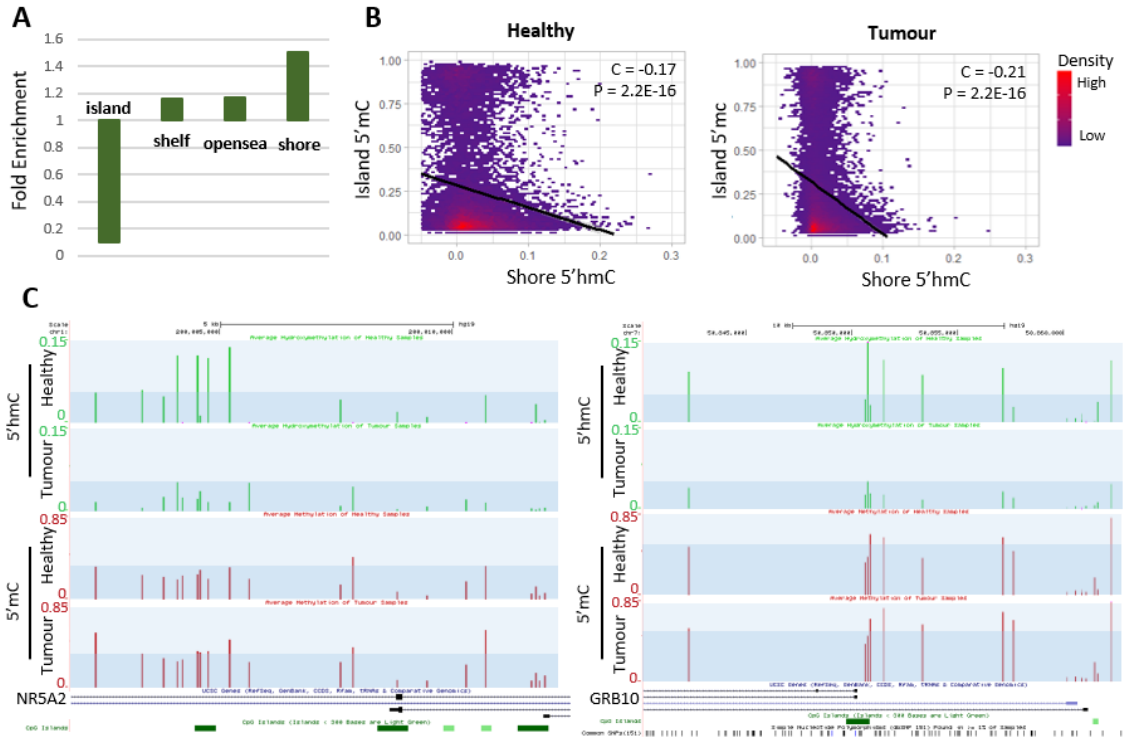


Figure 3.7: **Loss of 5'hmC at CpG Shores Correlates with increased 5'mC at CpG Islands**. **A**) Distribution of highly hydroxymethylated probes in relation to CpG islands. Hydroxymethylation was depleted at CpG islands and enriched at CpG shores. **B**) Density scatterplot showing correlation between average 5'mC at CpG islands and 5'hmC at each islands shore in healthy and tumour samples. **C**) Genome browser views of average healthy and tumour 5'mC (red) and 5'hmC (green) across the NR5A2 and GRB10 genes, which are silenced by methylation in PDAC.

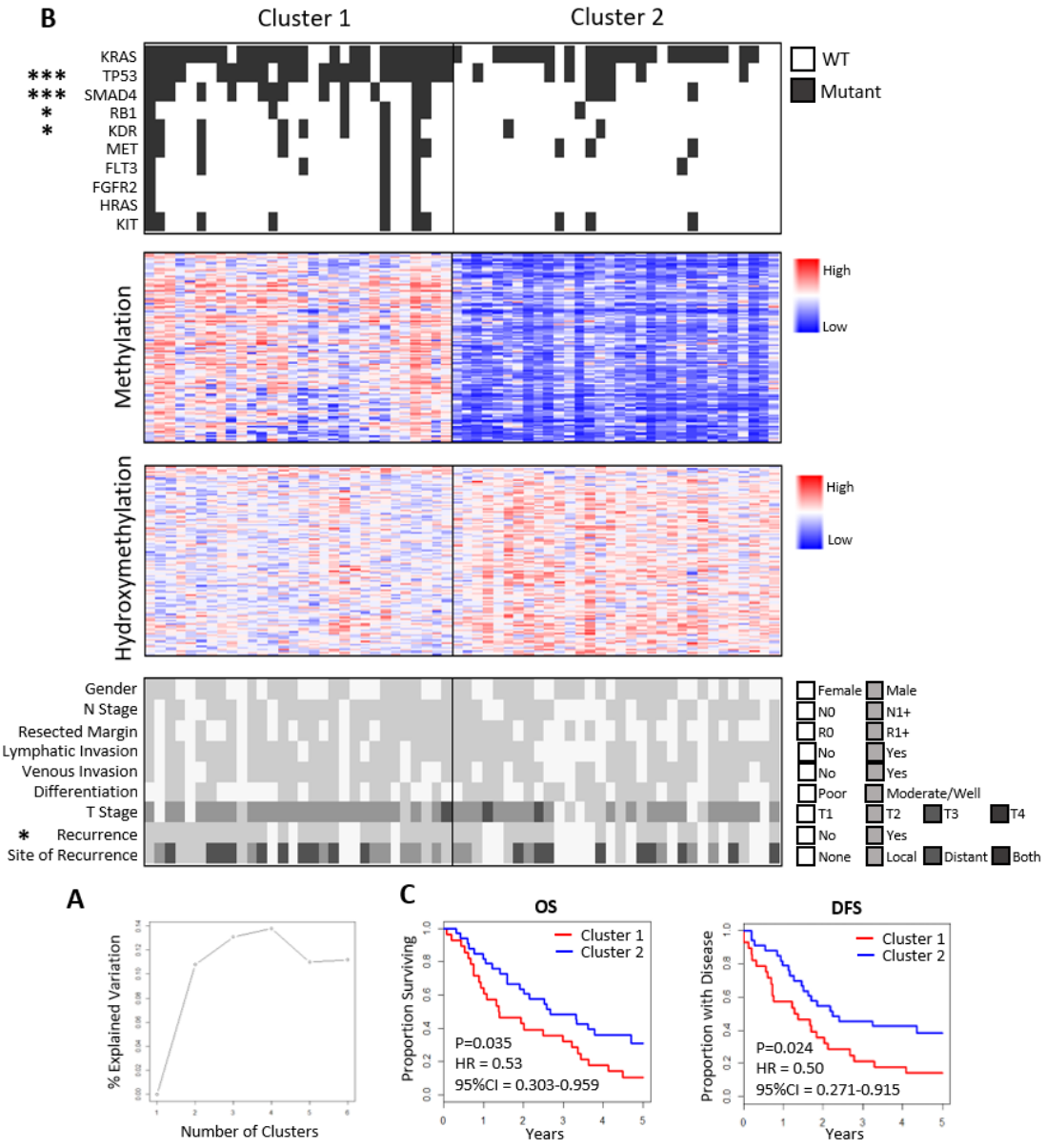


Figure 3.8: **Integrated Genetic and Epigenetic Analysis of PDAC**. **A**) BIC results showing percentage of variation explained for each number of clusters. **B**) iCluster analysis of pancreatic tumours reveals two molecular subtypes defined by their genetic and epigenetic profiles. Mutations are shown in the top panel. Bottom panel shows clinical characteristics. Genes and clinical features significantly associated with a cluster are indicated with an *. **B**) Kaplan-Meier curves for clusters. Patients in Cluster 1 had significantly worse overall (OS) and disease free survival (DFS).

3.2.6 Integrated Molecular Analysis of PDAC Links 5'hmC with PDAC Molecular Subtype

In order to see how 5'hmC fits in with the greater genetic and epigenetic landscape, integrated molecular analysis was performed using the programme iCluster (Shen et al., 2009), which can simultaneously cluster multiple data types. The Top 5000 most variable methylated and hydroxymethylated probes were used, along with binary mutation data for any genes mutated in more than 5% of patients. Bayesian information criteria (BIC) modelling using the iCluster programme showed that most of the variation between samples could be explained with two clusters (Fig. 3.8A). These clusters were approximately equally sized and distinct in their genetic, epigenetic and clinical profiles. Cluster 1 was characterized by low levels of 5'hmC and increased hypermethylation of a number of probes, as well as an increased incidence of TP53, SMAD4, RB Transcriptional Corepressor 1 (RB1) and Kinase Insert Domain Receptor (KDR) mutations. Cluster 2 on the other hand had higher 5'hmC levels, more hypomethylated probes and fewer mutations (Fig. 3.8B). Tumours in Cluster 1 were more aggressive, with overall and disease free survival approximately half that of patients in Cluster 2 (Hazard Ratio of 0.53 and 0.50 respectively) (Fig. 3.8C). Cluster 1 was also associated with an increased rate of recurrence (Fig. 3.8A), but not with any other clinical characteristics, overall indicating a more aggressive set of tumours.

mRNA expression data was available for 46 patients in the iCluster analysis. Gene set enrichment analyses of diferentially expressed transcriptional programmes revealed increased activation of signalling pathways known to promote PDAC tumour progression in Cluster 1, including TGFβ, JAK STAT and Myc targets. Additionally, gene sets linked to EMT and tumour metastasis (Zeb1 targets, cell junction organization, Liver metastasis up), as well as increased angiogenesis and Cisplatin resistance indicate that overall Cluster 1 tumours are more aggressive. Interestingly, a number of gene sets enriched in Cluster 1 have also been linked to the aggressive

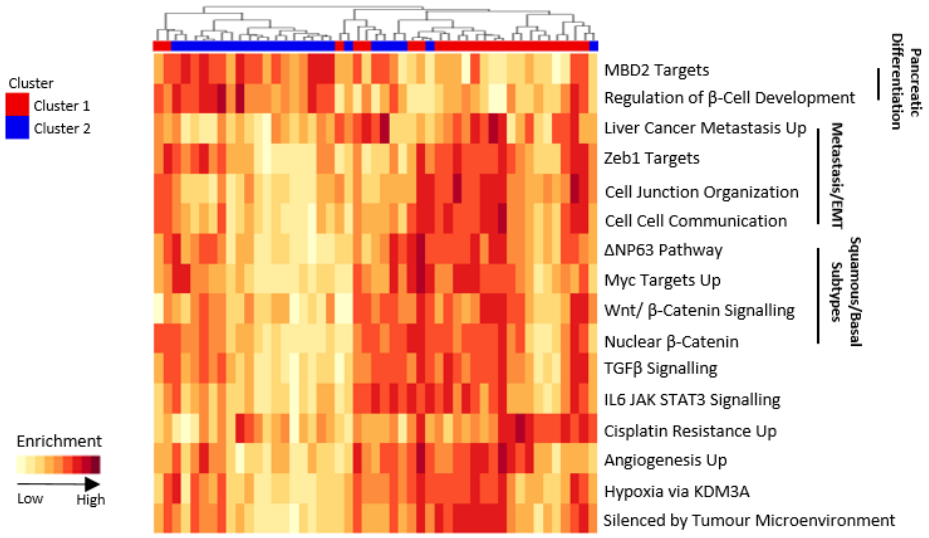


Figure 3.9: **Differentially Expressed Gene Sets Between Clusters**. **A**) Significant differentially expressed gene sets between Cluster 1 and Cluster 2 Tumours. Cluster 1 is enriched for gene sets related to metastasis and signalling pathways involved in more aggressive molecular subtypes, while Cluster 2 is enriched for gene sets related to pancreatic differentiation.

squamous molecular subtypes, such as ΔNP63, Myc and Wnt pathway signalling (Bailey et al., 2016)(Seino et al., 2018). Furthermore, Cluster 2 was enriched for gene sets related to terminal pancreatic differentiation, such as β-cell development. Notably, Cluster 2 was enriched for Methyl-CpG Binding Domain Protein 2 (MBD2) targets, which includes a number of genes important for exocrine pancreas differentiation (Berger J, 2007), many of which are associated with one or more less aggressive (classical-pancreatic) subtypes, including CEL, PRSS2, REG1B and TFF2 (Bailey et al., 2016)(Moffitt et al., 2015)(Collisson et al., 2011). In order to see how these subtypes correspond with the genetic/epigenetic clusters identified here, tumours were classified according the gene sets described by Bailey, Moffitt and Collisson using mRNA expression data. A gene set was created for each subtype and the relative enrichment of each gene set across the whole cohort was determined using the programme GSVA. Tumours were then hierarchically clustered and assigned subtypes according to the most enriched gene set in each cluster. (Fig. 3.10A; Appendix). In

all cases, the aggressive subtype (Squamous for Bailey, Quasimesenchymal [QM PDA] for Collisson and Basal for Moffitt) had a worse overall survival (Appendix).

There was extensive overlap between aggressive subtypes; consistent with recent proposals that each signature is identifying a similar set of aggressive tumours (Collisson et al., 2019). Because of this any tumours in the Oxford classified as squamous, basal or QM-PDA were grouped together into a single subtype termed squamous and the remainder classified as classical-pancreatic (Fig. 3.10A, top panel). The squamous subtype was more aggressive than classical-pancreatic as patients with classical-pancreatic tumours had survival rates over double that of squamous (Fig. 3.10B). Interestingly, over 95% of tumours in Cluster 1 from the iCluster analysis were classified as belonging to this squamous subtype (Fig. 3.10C). Cluster 2 on the other hand was composed mostly of Classical tumours. Cluster 1 may therefore reflect the genetic and epigenetic underpinnings of a subset of Aggressive tumours.

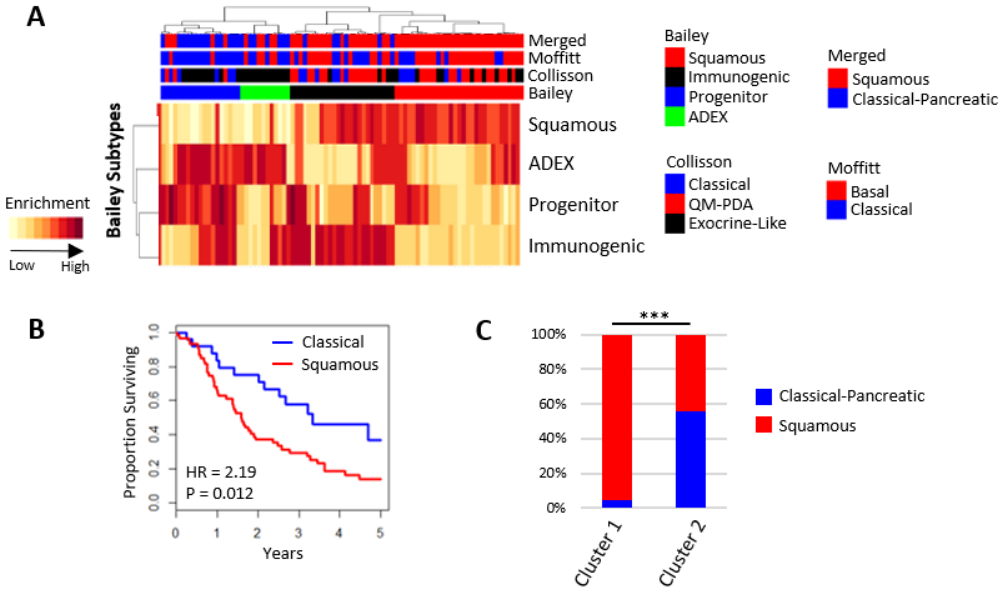


Figure 3.10: **Molecular Subtyping of the Oxford Cohort**. **A**) Molecular subtyping of tumours using gene sets defined by Bailey et al. Relative enrichment of each gene set was determined using GSVA. The determined subtype for Bailey, Collisson and Moffitt are shown in the above panel as well as the merged subtypes. **B**) Kaplan-Meier curves for patients with squamous and classical-pancreatic tumours. **C**) Distribution of squamous and classical-pancreatic subtypes in Clusters identified in figure 1.8. Squamous tumours are enriched in Cluster 1.

To validate some of the results seen in the Oxford cohort and to determine if these clusters are reliable indicators of patient outcome, an epigenetic signature was derived for each cluster and applied to the TCGA PAAD cohort. 5'mC data was used for this as 5'hmC information is not available in any other PDAC cohort. Initially, differentially methylated probe (DMP) analysis was performed between clusters. Gene set enrichment of DMPs showed that Cluster 1 tumours were hypermethylated for Polycomb repressor complex (PRC) targets, as well as for genes silenced by trimethylation of histone 3 lysine 27 (H3K27me3) in induced pluripotent stem (IPS) cells, overall indicating a more stem-like methylome (Fig. 3.11A). The most highly enriched gene set was for members of the Hox family; a phenomenon frequently reported in pancreatic cancer that is usually attributed to the developmental role of the Hox genes during pancreatic differentiation (Rodrigues et al., 2016)(Gray et al., 2011). Interestingly, the HOXA cluster is developmentally regulated by TET2, the main 5'hmC enzyme in the adult pancreas, further linking low 5'hmC levels with methylation of tumour suppressors (Bocker et al., 2012). A predictive subclass probeset was created using the top 200 differentially methylated probes between each cluster. When applied to the PAAD TCGA dataset, patients also separated into two distinct groups (Fig. 3.11B), with Cluster 1 methylome patients having an overall survival half of that of Cluster 2 patients (HR = 0.50) (Fig. 3.11C). Furthermore, Cluster 1 patients in this cohort were enriched for squamous tumours and highly enriched for SMAD4 mutations, but not TP53, RB1 and KDR mutations. Therefore, the epigenetic signature derived from multi-omic clustering of patients is a robust predictor of patient survival and molecular identity in both tested datasets.

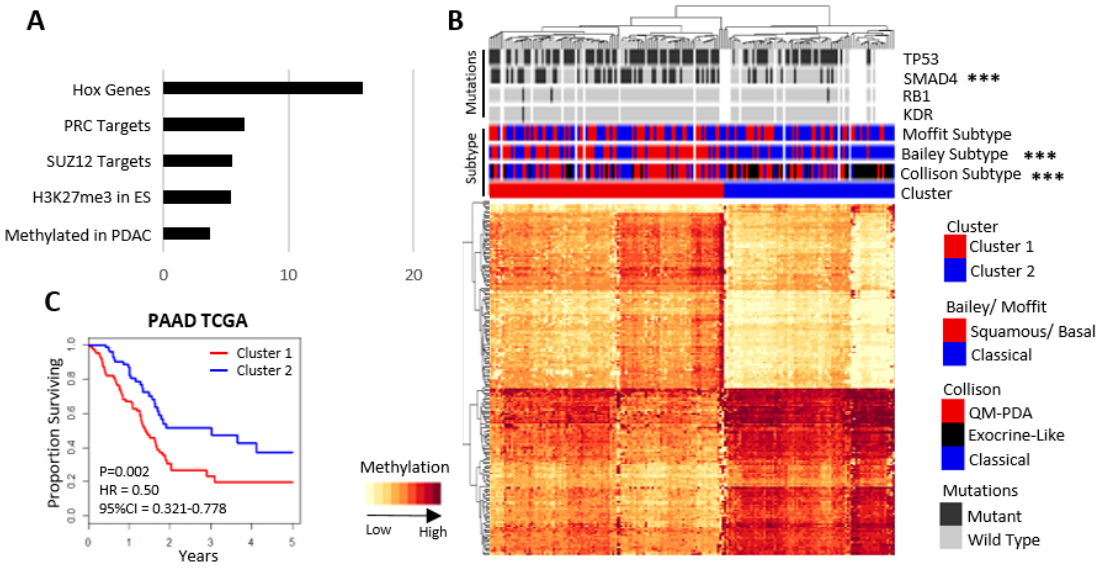


Figure 3.11: **Application of Cluster Methylation Signature to another Cohort**. **A**) Most significant differentially methylated gene sets overrepresented in Cluster 1. The Hox genes were highly methylated in Cluster 1, as well as Polycomb Repressor Complex (PRC) targets. **B**) Classification of TCGA Pancreatic Adenocarcinoma (PAAD) dataset using top differentially methylated probes shows clear separation into two clusters. Mutations in genes that were significantly associated with a cluster in the Oxford cohort are shown above, as well as the molecular subtype of each tumour. SMAD4 mutations and squamous subtypes were enriched in Cluster 1 in both cohorts. **C**) Kaplan-Meier curve for clusters in TCGA cohort showing decreased survival of patients in Cluster 1 methylome patients.

Given the association between low 5'hmC and Cluster 1, the relationship between 5'hmC and molecular subtypes was investigated further. Interestingly, the patients with the lowest levels of 5'hmC across tumour specific DMhPs identified in Figure 3.6A were almost entirely made up of the Aggressive subtype (Fig. 3.12A). Loss of the endodermal differentiation marker GATA6 has been linked to all three aggressive subtypes described by Collisson, Bailey and Moffitt, and squamous tumours indeed have lower levels of GATA6 mRNA in the Oxford cohort (Fig. 3.12B). Examination of the distribution of 5'hmC across the GATA6 gene in classical-pancreatic and squamous tumours showed a drop in 5'hmC in squamous tumours at the distal end of the gene (probe cg21040686) (Fig. 3.12C,D). Furthermore, the level of 5'hmC at this probe correlated with the extent of 5'mC at the GATA6 promoter CpG island

(mean β-value across all probes within the CpG island) (Fig. 3.12D) and positively correlated with GATA6 mRNA (Fig. 3.12E), suggesting that drop in 5'hmC at the GATA6 gene coincides with its epigenetic silencing. These results link 5'hmC with tumour differentiation and subtype identity, with the lowest levels of 5'hmC associated with aggressive, poorly differentiated tumours and 5'hmC loss occurring at important classical-pancreatic subtype genes.

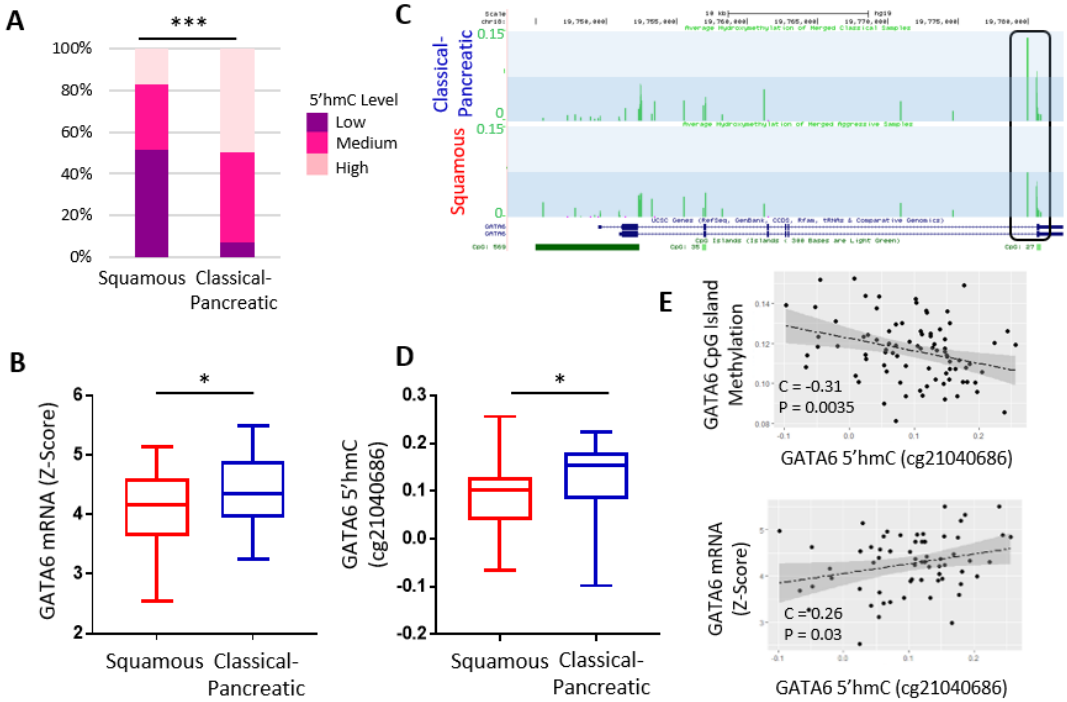


Figure 3.12: **5'hmC Correlates with PDAC Molecular Subtype**. **A**) Distribution of patients with low, medium or high 5'hmC identified in figure 1.6 in squamous and classical-pancreatic subtype tumours. Patients with Low 5'hmC were almost entirely composed of squamous tumours. **B**) Differences in GATA6 mRNA between subtypes. Classical-pancreatic tumours have higher GATA6 mRNA. **C**) Genome track showing average 5'hmC across the GATA6 gene in squamous and classical-pancreatic tumours. A drop in 5'hmC is seen at the distal end of the gene in squamous tumours (black box). **D**) 5'hmC level at this probe in different subtypes. **E**) Correlation of 5'hmC at probe cg21040686 and GATA6 promoter CpG island methylation and GATA6 mRNA. Loss of 5'hmC at the GATA6 gene correlates with increased promoter methylation and decreased gene expression

3.3 Discussion

In this chapter, the role of the hydroxymethylome was investigated in a cohort of PDAC patients. Initial experiments using IHC to measure total 5'hmC in tumours revealed that 5'hmC is progressively lost at a global scale during PDAC tumorigenesis. Genome wide 5'hmC profiling was then performed in order to determine where in the genome 5'hmC changes were occurring. In order to accomplish this, the oxBS method was adapted for use with FFPE material for the first time. There are a number of pieces of evidence that indicate that adaption of this method was successful. Firstly, raw density plots from the EPIC arrays showed a drop in signal following oxidative treatment due to the conversion of 5'hmC to 5'fC. Additionally, when subtracting the oxBS signal from BS, there were far more positive $\Delta\beta$-values (representing potential hydroxymethylation) than negative (representing background noise). The detected hydroxymethylation signal was enriched at CpG shores and depleted a CpG islands; a pattern described previously (Stewart et al., 2015), and there was a drop in hydroxymethylation in tumour samples relative to healthy tissue, that is consistent with IHC results and has been reported for various cancers (Qiu et al., 2018)(Zhong et al., 2018). Finally, tumour specific hydroxymethylation was lost at genes that are specifically downregulated in pancreatic cancer. This is a strong indication that this signal is real and not just background noise. It should be noted that the 5'hmC signal detected was far weaker than 5'mC. However, 5'hmC is a far less abundant epigenetic mark (Tahiliani et al., 2009)(Ito et al., 2011)(Thomson et al., 2013)(Globisch et al., 2010) and the 5'hmC values here are similar to other studies using methylation arrays in different tumour types (Johnson et al., 2016). Additionally, EPIC array probes are highly enriched at CpG islands due to their role as regulatory elements, which is where 5'hmC is specifically depleted. Taken together, these data show that it is possible to generate genome wide 5'hmC data using FFPE material; the most common method for storage of patient tissue. This opens up this method for

use with a wealth of biobanked samples that will greatly enhance the understanding of 5'hmC in the genome.

Genome wide 5'hmC profiling in PDAC tumours showed a decrease in 5'hmC at genes critical for PDAC development, including the MAPK pathway and TP53 targets; two of the main drivers of pancreatic cancer. This reduction in global 5'hmC during tumorigenesis is reminiscent of results by Li et al. (2018) who quantified changes in global 5'hmC levels by dot blot when differentiating embryonic stem (ES) cells to pancreatic epithelial (PE) cells. Both ES and PE have high levels of 5'hmC, but the intermediate stages between these two defined cell types have more than a 50% drop in 5'hmC levels. Therefore, the drop in global 5'hmC may reflect a reversion of differentiated pancreatic cells to a more progenitor-like state as they become neoplastic. This idea is supported by the fact that PDAC tumours are losing 5'hmC at genes involved in the regulation of the insulin pathway. Furthermore, multi-omic analysis showed that tumours with the lowest levels of 5'hmC are associated with a more stem-like methylome, as well as decreased expression of genes involved in exocrine pancreas differentiation. In ES cells, 5'hmC defines a set of pluripotency associated genes and loss of 5'hmC results in spontaneous differentiation due to loss of control of these genes (Ito et al., 2010)(Tovy et al., 2017). In differentiated pancreatic cells, where 5'hmC instead maintains a number of pancreatic differentiation genes (Li et al., 2018), it follows that a reduction in 5'hmC would result in loss of control of these genes, resulting in a reversion to a less differentiated, and more progenitor-like, state.

Interestingly, lower 5'hmC levels were also associated with more aggressive subtypes of pancreatic cancer. Furthermore, at the GATA6 gene, there was a decrease in 5'hmC in squamous tumours which correlated with decreased gene expression and increased promoter methylation, suggesting that GATA6 can be regulated by 5'hmC. GATA6 is a zinc finger transcription factor critical for pancreatic differentiation (Tiyaboonchai et al., 2017)(Xuan and Sussel, 2016) that is highly expressed

in classical-pancreatic tumours but weakly expressed in squamous subtypes. In addition, GATA6 expression alone was found to be a robust surrogate biomarker for distinguishing between these subtypes (Aung et al., 2018), suggesting GATA6 itself is regulating important targets for maintaining a classical identity, or GATA6 is being co-regulated with other genes that are doing this. Therefore, 5'hmC may be important for maintaining a classical-pancreatic subtype identity by regulating classical associated genes such as GATA6. The potential link between loss of 5'hmc, GATA6 and Aggressive molecular subtypes presented in this chapter warrants further study. However, these results are all correlative and so it remains unclear is 5'hmC itself is actively regulating the expression of these genes, or is merely a carrier mark for their activation. This question will form the focus of the next chapter.

Chapter 4

TET2 Promotes GATA6 and can be Targeted Pharmacologically

4.1 Introduction

The work is the previous chapter suggested a link between loss of 5'hmC and GATA6, as well as progression to a more aggressive phenotype. However, genomic studies such as these can only draw correlative conclusions, and cannot indicate if loss of 5'hmC is causative or not. Therefore, in the following chapter the causative effects of modulating 5'hmC levels will be examined, with a focus on TET2, the main 5'hmC enzyme in the adult pancreas (Li et al., 2018). This will be done using overexpression techniques as well as pharmacologically, using drugs that alter the stability and activity of TET2: Metformin and Vitamin C.

4.1.1 Vitamin C (Ascorbic Acid) Potentates TET Activity

TET enzymes are members of the 2-OG dioxygenase family, and require 2-OG and Fe^{2+} as co-factors for their catalytic activity (Kuiper and Vissers, 2014). Vitamin C (Also known as ascorbic acid) is a well characterized anti-oxidant that can reduce

Fe^{3+} to Fe^{2+}, thereby enhancing the catalytic activity of TET enzymes (Hore, 2017). Additionally, Ascorbic Acid (AA) can directly interact with the C-terminal catalytic domain of TETs, possibly causing conformational changes (Yin et al., 2013). This direct interaction may explain why ascorbic acid is much more efficient in enhancing TET activity compared to other antioxidants (Hore et al., 2016). At low concentrations ($\sim$100μM-1mM) AA has been shown to enhance TET activity and increase 5'hmC levels both *in vitro* and *in vivo*, leading to widespread, but specific, DNA demethylation (Yin et al., 2013)(Wu and Zhang, 2014)(Liu et al., 2019).

Physiologically, AA levels can range from 0-0.4mM (Zhang et al., 2002). However, intravenous administration can achieve concentrations over 8mM (Chen et al., 2007). At these concentrations high levels of hydrogen peroxide (H202) are generated, and AA switches from being an anti-oxidant to an oxidant. At these doses it is likely that AA no longer enhances TET activity, or could suppress it. However, AA can be sub compartmentalized in cells (D'Aniello et al., 2017), so it is difficult to determine if this is indeed the case. High dose intravenous AA has been shown to suppress the growth and metastasis of pancreatic tumours in mice (Polireddy et al., 2017), as high levels of H2O2 can selectively kill cancer cells (Chen et al., 2005)(Chen et al., 2007), though clinical trials in humans have so far been small and inconclusive (Polireddy et al., 2017)(Drisko et al., 2018)(Gonzalez M, 2016)(Welsh et al., 2013). Interestingly, the anti-metastatic effects of ascorbic acid described by Polireddy et al. (2017) were also present at lower, sub-cytotoxic concentrations ($\sim$1mM), suggesting that this effect may be due to modulation of 5'hmC rather than generation of H202.

4.1.2 Regulation of TET2 by AMPK and Metformin

AMP-activated protein kinase (AMPK) plays a crucial role in maintaining cellular energy homeostasis in response to intracellular ATP levels. Under low ATP levels, such as starvation or hypoglycemia, AMPK becomes phosphorylated at Thr172 (pAMPK)

which causes it to switch on catabolic pathways to generate ATP, and switch off anabolic pathways that consume ATP (Hardie, 2011). AMPK has been proposed as a tumour suppressor in PDAC, and its loss can promote cell metastasis (Chen et al., 2017), and induce resistance to Gemcitabine treatment (Zhu et al., 2018). Critically, AMPK is suppressed in patients with diabetes or hyperglycemia (Coughlan et al., 2014), which is found in approximately 80% of PDAC patients at the time of diagnosis (Permert et al., 1993)(Schwarts et al., 1978), many of which have no previous history of diabetes (Mekan et al., 2006). Wu et al. (2018) identified TET2 as a substrate of AMPK, and AMPK-induced phosporylation at Ser99 promotes TET2 protein stability. Thus under the hyperglycaemic conditions associated with PDAC, TET2 becomes destabilized and 5'hmC levels decrease.

The anti-diabetic drug Metformin has long been linked to preventing the development of PDAC, and is associated with increased survival in diabetic patients with pancreatic cancer (Wang et al., 2014)(Sadeghi et al., 2012). One mechanism of action for Metformin is to inhibit ATP production by regulation of the electron transport chain, leading to activation of AMPK (Evans et al., 2005)(El-Mir et al., 2000). Therefore, Metformin can be used in order to stabilize TET2 protein levels to reinstate 5'hmC levels (Wu et al., 2018).

4.1.3 Aims

In the following chapter, the causative effects of TET2 and 5'hmC on the expression of GATA6, as well as other markers of molecular subtype will be examined in pancreatic cell lines. TET2 abundance will be altered by overexpression experiments and TET2 stability and activity will be modulated using Metformin and ascorbic acid. These experiments will determine if 5'hmC is playing an active role in regulating these genes, or is simply a marker of their activation.

4.2 Results

4.2.1 Molecular Classification of Pancreatic Cell Lines

In order to determine the effects of 5'hmC on molecular subtype markers in pancreatic cell lines, the subtype of these cell lines must first be known. Gene expression data for 44 PDAC cell lines was available from The Cancer Cell Line Encyclopedia (CCLE) (Barretina et al., 2012). The relative enrichment of subtype gene sets was then determined using GSVA. The exocrine-like and ADEX subtypes were not included as they have not been found to be represented in PDAC cell lines, and the immunogenic subtype was excluded as it is dependent on immune populations in the tumour microenvironement that will not be present in tumour derived cell lines. This left six subtype gene sets: three of which of which were associated with the squamous subtype (squamous [Bailey], basal [Moffit] and QM PDA [Collisson]), and three which were associated with the classical-pancreatic subtype (progenitor [Bailey], classical [Moffitt] and classical [Collisson]). Hierarchical clustering of enrichment scores resulted in two clusters which were enriched for either squamous or classical-pancreatic subtypes, and cell lines were classified accordingly (Fig. 4.1A). These results were compared to those of Collisson et al. (2011), who previously classified 19 human PDAC cell lines as either Classical or QM-PDA, 15 of which were also present in the CCLE dataset. Of these cell lines, 14 were classified as their corresponding squamous/classical-pancreatic subtype, while one cell line (HUPT3) was not. Cell lines classified as squamous had a significantly lower level of GATA6 than classical-pancreatic cell lines, suggesting correct classification of these cells (Fig. 4.1B). Of the classified cells, 3 squamous cell lines (PANC1, PSN1 and MiaPaca2) and 1 classical-pancreatic cell line (CAPAN1) were selected.

There are three members of the TET family. TET1 is the most highly expressed in embryonic stem cells, definitive endoderm and gut tube cells, while TET2 and TET3

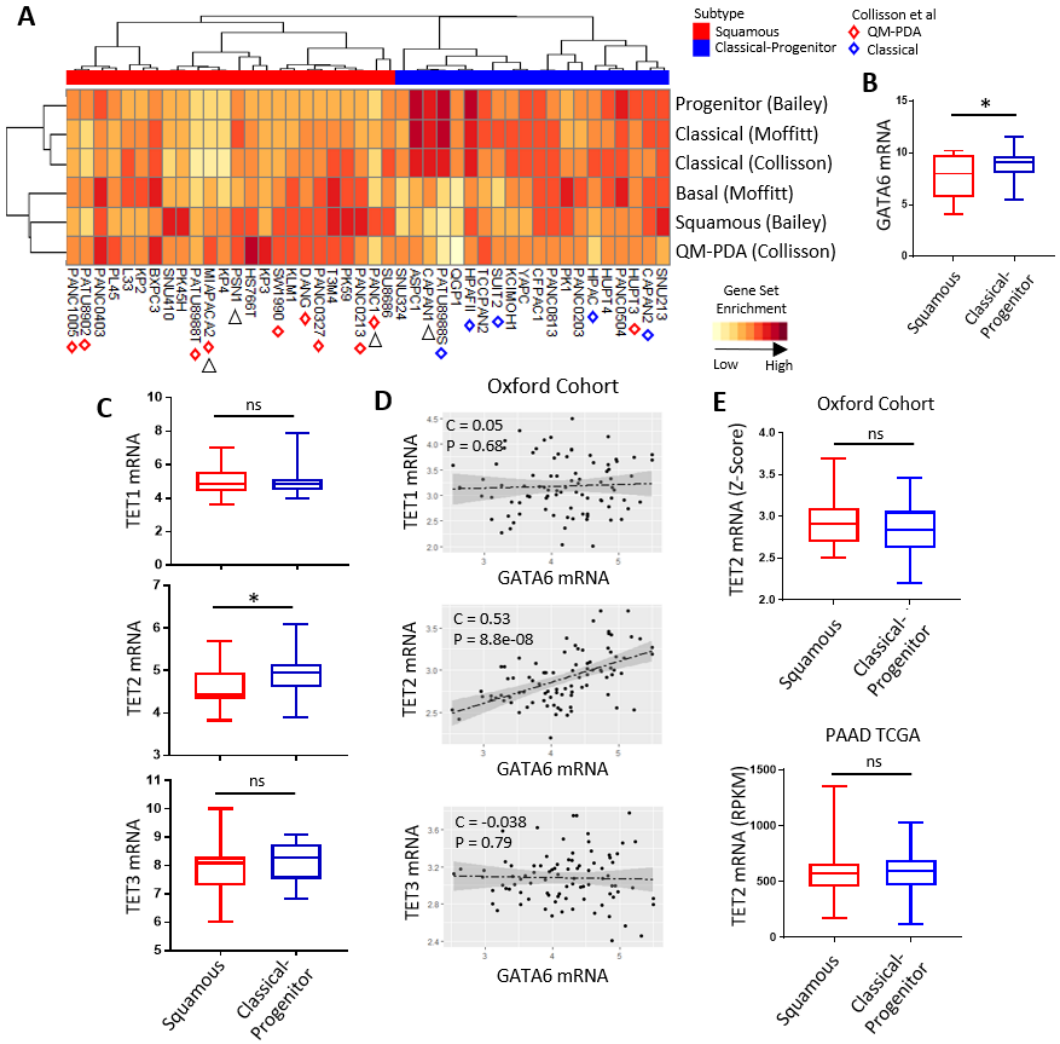


Figure 4.1: **Molecular Classification of PDAC Cell Lines**. **A**) Heatmap showing hierarchically clustered enrichments for subtype specific gene sets. Cell lines split into two groups that were enriched for either squamous or classical-pancreatic subtype gene sets. Red/blue triangles indicate cell lines classified by Collisson et al. (2011) in supplementary figure 2. Empty arrowheads indicate cell lines used for experiments. **B**) Box and whisker plot showing differences in TET1, TET2 and TET3 mRNA for squamous and classical-pancreatic cell lines from CCLE data. Squamous cell lines have lower levels of TET2. **C**) Correlations between TET1, TET2 and TET3 with GATA6 mRNA in the Oxford cohort. GATA6 correlates with TET2, but not TET1 or TET3. **(D)** Differences in TET2 mRNA between squamous and classical tumours in the Oxford and PAAD TCGA cohorts.

are highly expressed in foregut cells. Only TET2 is highly expressed in the adult pancreas (Li et al., 2018). In PDAC cell lines, squamous cells had significantly lower levels of TET2 mRNA than classical-pancreatic cells, while there was no difference in TET1 and TET3 levels (Fig. 4.1C). Additionally, in the Oxford cohort there was a significant positive correlation between TET2 and GATA6 mRNA, but not between GATA6 and TET1 or TET3 (Fig. 4.1D), suggesting that TET2 is the main TET enzyme associated with molecular subtypes in the pancreas. However, there was no difference in TET2 mRNA levels in squamous and classical-pancreatic tumour in the Oxford or PAAD TCGA cohorts (Fig. 4.1E), possibly due to the fact that these cohorts are whole tissues composed of a mix of tumour, stroma and healthy cells while PDAC cell lines are only tumour cells.

4.2.2 Overexpression of TET2 in Aggressive Cells Increases Classical-Pancreatic Subtype Markers

To see if TET2 can actively regulate molecular subtype genes, FLAG-TET2 was overexpressed in PANC1 and PSN1 cells and the expression of subtype markers was measured by western blot. Loss of endodermal differentiation markers such as GATA6 and PDX1 is characteristic of squamous tumours, with low GATA6 expression being the most reliable single marker of squamous subtypes found so far (Aung et al., 2018). Additionally, classical subtypes tend to be more epithelial, while squamous subtypes are more mesenchymal. Therefore, E-Cadherin and N-Cadherin were also chosen as markers for classical and squamous subtypes respectively. In CCLE cell lines, squamous cells were associated with significantly lower levels of GATA6 and E-Cadherin mRNA, but not PDX1 or N-Cadherin (Fig. 4.1B/ Appendix). In previous studies TP63 isoform expression has been used as a marker for molecular subtype, as the oncogenic ΔN-p63 isoform is highly expressed in the squamous tumours (Bailey et al., 2016). However, in a panel of 6 Aggressive cell lines, including PANC1 and

MiaPaca2, ΔN-p63 was expressed in only 1 cell line (BxPC3). The remaining cell lines were either completely negative for TP63 at both the mRNA and protein level, or expressed the TA-p63 isoform (Andricovich et al., 2018), suggesting that this is not a suitable marker for subtypes in cell lines. Therefore, TP63 expression was not used in this study. TET2 overexpression successfully increased global 5'hmC levels by dot blot (Fig. 4.2A), and TET2 protein levels by western blot (Fig. 4.2B). PDX1 and GATA6 levels were also increased in both cell lines, suggesting increased differentiation of cells upon TET2 expression. Additionally, TET2 increased E-Cadherin levels in both cell lines, and decreased N-Cadherin in PANC1 cells. N-Cadherin was absent in control PSN1 cells and this did not change upon TET2 overexpression. These changes are consistent with a switch in molecular subtype of cells, suggesting that these squamous cell lines are becoming more classical in identity, and indicating that TET2 can actively regulate subtype specific genes.

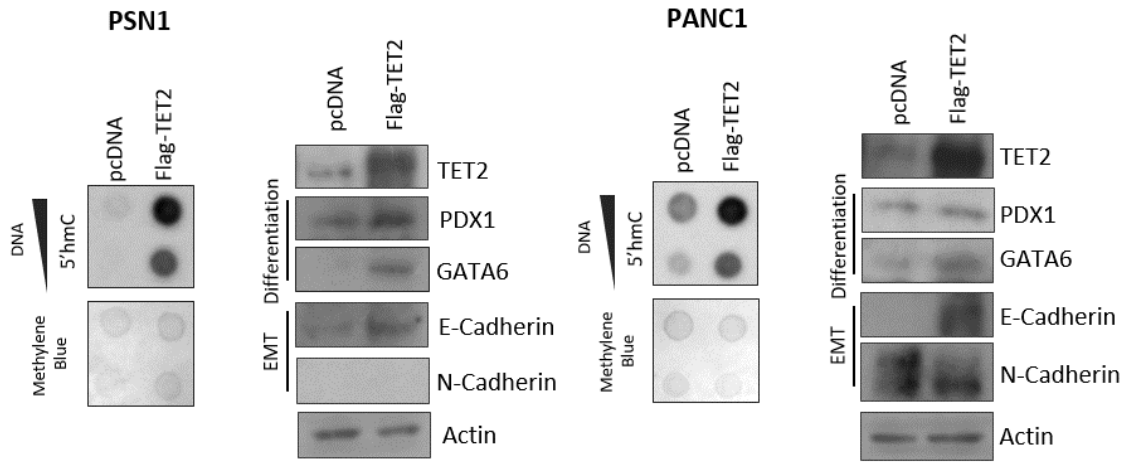


Figure 4.2: **Overexpression of TET2 in Aggressive Cell Lines**. TET2 was overexpressed in (**A**) PSN1 and (**B**) PANC1 cells. Left) Global 5'hmC levels were measured by dot blot, and (right) protein levels determined by western blot. TET2 overexpression increased 5'hmC and TET2 levels. GATA6, PDX1 and E-Cadherin, markers of Classical subtypes, were also increased. N-Cadherin, a marker of Aggressive subtypes, was decreased in PANC1 cells.

4.2.3 Metformin can increase TET2 and GATA6 levels in an AMPK-dependent manner

While TET2 expression is decreased in squamous cell lines, the TET2 gene is rarely mutated in PDAC. Within the PAAD TCGA cohort, TET2 somatic mutations were present in only 4 out of 149 PDAC patients (2.7%). Similarly, TET2 is mutated in only 2 of 41 PDAC cell lines from the CCLE (4.8%) (Appendix). TET2 was not present on the Hot Spot array employed to genetically characterize the Oxford Cohort. Therefore, in the majority of patients TET2 must be downregulated through some other mechanism, such as protein stability, raising the possibility that TET2 levels could be increased pharmacologically. Phosphorylation of TET2 by AMPK results in its stabilization, which is inhibited under hyperglycaemic conditions such as diabetes. The anti-diabetic drug Metformin can restore AMPK activity, resulting in re-stabilization of TET2 (Wu et al., 2018). Therefore, treatment of squamous cells with Metformin may increase GATA6 expression by stabilizing TET2 level (Fig. 4.3A).

In order to test this hypothesis, PSN1 cells were treated with Metformin for 48 hours, with fresh media containing the drug replaced every 24 hours. Patients taking Metformin have a circulating concentration of $\sim$20μM (Dowling et al., 2012)(Barbieri et al., 2015). However, typically *in vitro* studies such as the one carried out by Wu et al. (2018) use concentrations far higher than this, ranging up to 5mM. In an effort to make results more applicable to clinical conditions concentrations were kept at low levels. Cells were treated with 20μM Metformin, or a higher dose of 500μM. Metformin successfully increased TET2 and total 5'hmC levels, as well as GATA6 protein levels in a dose dependent manner (Fig. 4.3B), indicating that Metformin can be used to make squamous cell lines more classical in identity.

Metformin has been shown to exert both AMPK-dependent and AMPK-independent effects (Rena et al., 2017). While not directly shown by Wu et al. (2018), the effect of Metformin on TET2 stability was proposed to be AMPK-dependent. To determine

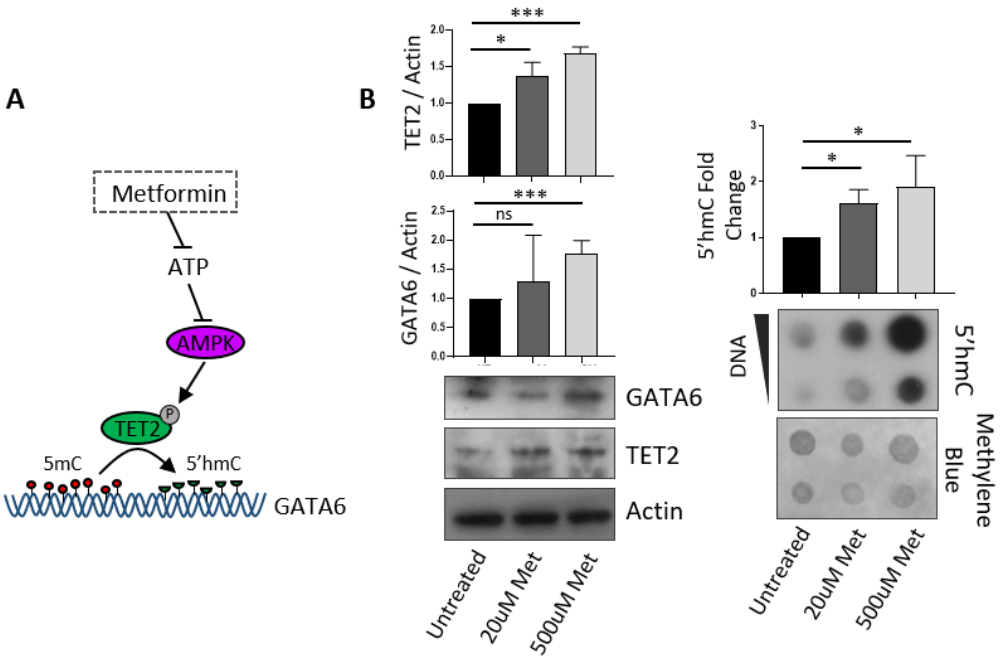


Figure 4.3: **Metformin Increases TET2 and GATA6 in PSN1 Cells**. **A**) Schematic showing action of Metformin on TET2. Metformin indirectly increases TET2 stability by regulating ATP levels. **B**) Western blot and **C**) dot blot for PSN1 cells treated with 20μM or 500μM for 48 hours. Metformin increased TET2 and GATA6 protein levels and global 5'hmC in a dose dependent manner

if this is indeed the case for the increases in TET2 and GATA6 seen here, the AMPK inhibitor Compound C was used. PSN1 cells were treated with Metformin in combination with either 20μM compound C in DMSO or DMSO alone as a negative control. Metformin in combination with DMSO acted in the same manner as Metformin alone. AMPK becomes activated by phosphorylation at threonine 172 (T172) in response to Metformin (Zhou et al., 2001). Compound C decreased AMPK phosphorylation and total AMPK levels, indicating successful AMPK inhibition. Treatment with compound C alone resulted in a decrease in TET2 levels, and addition of low or high dose Metformin was not able to rescue this. similarly, no increase in GATA6 was seen. This indicates that AMPK is acting downstream of Metformin, and the increases in TET2 and GATA6 seen are AMPK-dependent.

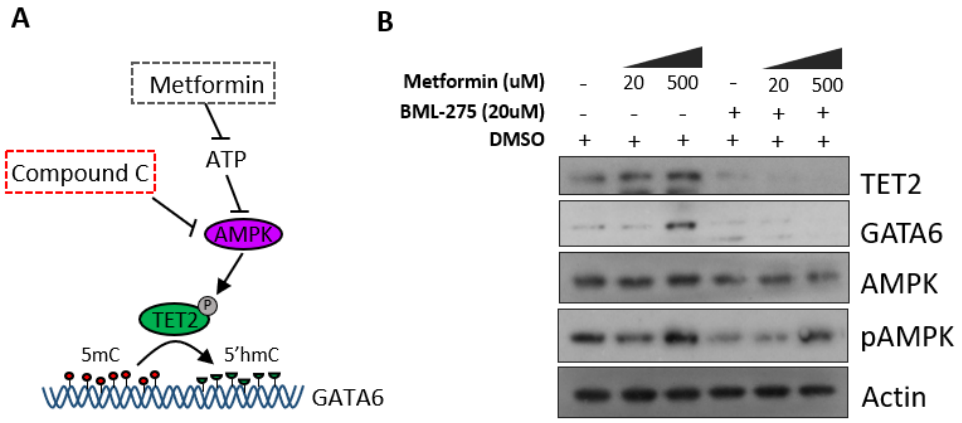


Figure 4.4: **Regulation of TET2 and GATA6 by Metformin is AMPK Dependent**. **A**) Diagram showing interaction between AMPK, Metformin and Compound C. **B**) PSN1 cells were treated with metformin and either Compound C in DMSO or DMSO only. Compound C inhibited AMPK activation and prevented metformin-induced TET2/GATA6 increases.

4.2.4 Restoration of 5'hmC by Ascorbic Acid Increases GATA6

Low dose ascorbic acid can enhance TET enzymatic activity and can mimic the effect of TET restoration (Cimmino et al., 2017). Thus AA treatment is another potential method to increase GATA6 expression in PDAC cells by increasing 5'hmC. PSN1 cells were treated with 100μM or 500μM AA for 48 hours, with fresh media containing the drug added every 24 hours. 100μM AA had little effect on global 5'hmC or GATA6 levels, however at 500μM a three-fold increase in global 5'hmC was seen, as well as a two-fold increase in GATA6 expression, indicating that low dose AA treatment can indeed increase GATA6 expression, but not at physiological concentrations.

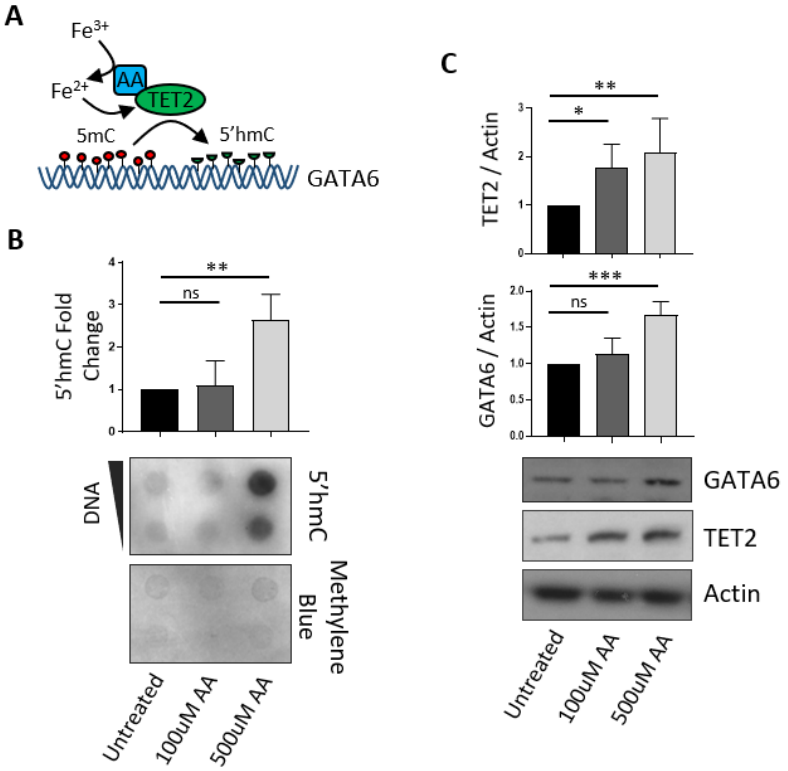


Figure 4.5: **Ascorbic Acid Increases 5'hmC and GATA6 Levels**. **A**) Schematic showing effect of ascorbic acid (AA) on TET2 activity. AA binds to TET2 and increases its catalytic activity by reducing Fe^{3+} to Fe^{2+}. **B**) Dot blot showing increase in 5'hmC in PSN1 cells treated with AA for 48 hours. **C**) Western blot showing increases in GATA6 and TET2 following AA treatment.

4.2.5 Metformin and Ascorbic Acid Act Synergistically to regulate 5'hmC and GATA6

Metformin and asorbic acid can both regulate TET2 by different mechanisms. However, clinically relevant metformin (20μM), and physiological ascorbic acid (100μM) had modest or no effect on GATA6 expression alone. Therefore, the combination of both drugs was tested in order to see if an increased effect could be seen by simultaneously increasing TET2 stability and activity.

PSN1 cells were treated for 48 hours with either 100μM ascorbic acid, 20μM metformin, or a combination of the two. AA alone induced a small, but significant, 10% increase in global 5'hmC levels, and metformin alone resulted in a 50% increase.

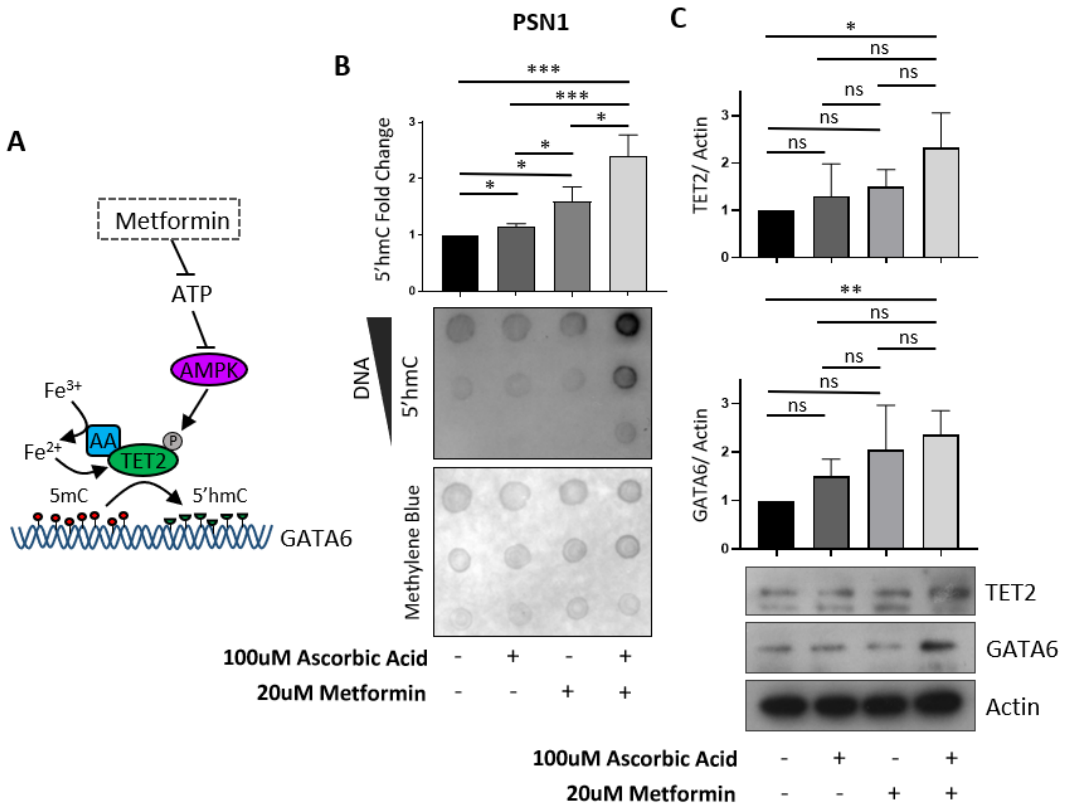


Figure 4.6: **Metformin and Ascorbic Acid Act Synergystically in PSN1 Cells**. **A**) Schematic showing interactions of Metfomin and Ascorbic Acid with TET2. **B**) PSN1 cells were treated with Ascorbic Acid or Metformin as single agents or combined. Combined treatment had a synergistic effect on total 5'hmC determined by dot blot and (**B**) increased TET2 and GATA6 protein levels by western blot.

Combined treatment increased 5'hmC 2.4 fold. This is higher than the sum of the effects of each drug individually indicating a synergistic relationship between the two drugs on 5'hmC levels. Individual treatment had little or no effect on both GATA6 and TET2 protein levels, but combined treatment could successfully increase TET2 and GATA6 2-3 fold. Therefore, the combination these drugs can be used *in vitro* at physiologically relevant levels to increase GATA6 expression and make squamous cells more classical-pancreatic in identity.

These results were then validated in two additional squamous cell lines. PANC1 and MiaPaca2 cells were treated for 96 hours rather than 48 due to their longer dou-

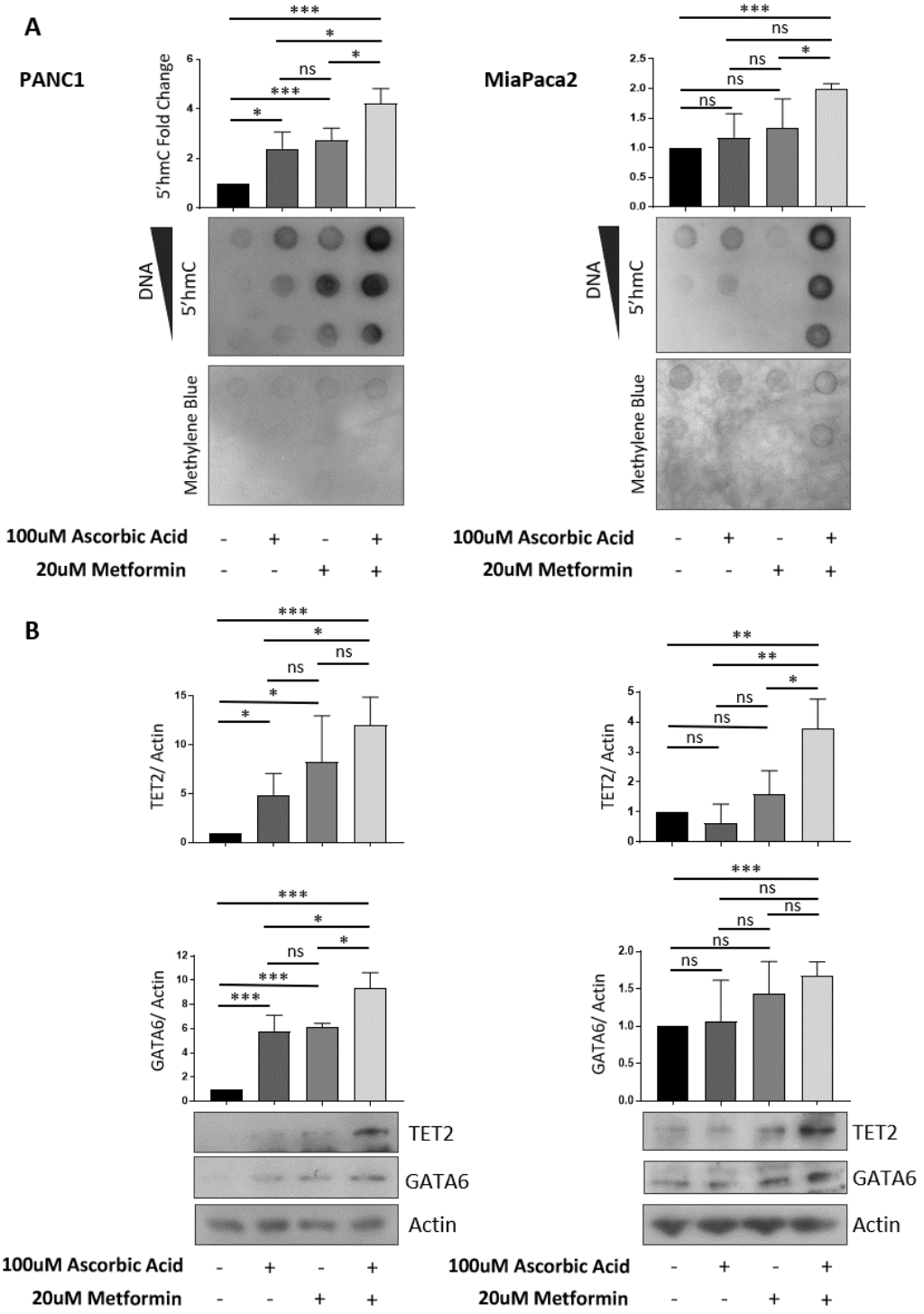


Figure 4.7: **Combined Metformin and Ascorbic Acid can increase GATA6 Expression in two additional Aggressive Cell Lines**. **A**) Dot blots from (left) PANC1 and (right) MiaPaca2 cells treated with Ascorbic Acid, Metformin, or a combination of the two. **B**) Western blots from treated cells. Combined treatment successfully increased TET2 and GATA6 expression in both cell lines.

bling time. In both cases, combined Metformin/AA treatment had an enhanced effect on 5'hmC levels than using single agents alone, and was successfully able to significantly increase GATA6 expression. In PANC1 and MiaPaca2 cells combined treatment had an additive effect on GATA6 levels. Therefore, combined Metformin/AA treatment represents a potential strategy to restore GATA6 expression in squamous cell lines in order to make them more classical in identity.

Next, the effect of these drugs on the Classical cell line was examined. CAPAN1 cells were treated for 96 hours with Metformin and AA either as single agents or in combination. Interestingly, AA alone increased global 5'hmC levels four-fold, and doubled TET2 protein levels, while metformin alone had little or no effect. Additionally AA and metformin combined had only a minor benefit to 5'hmC and TET2 levels compared to AA alone, suggesting that metformin is not regulating TET2 protein levels in this cell line, at least at the concentration used. This may be due to the fact that CAPAN1 cells have high levels of pAMPK relative to most other PDAC cell lines (Duong et al., 2012) so TET2 is already being stabilized by AMPK. Endogenous levels of GATA6 were high in CAPAN1 cells but could be further enhanced by treatment with metformin and AA, indicating that increases in 5'hmC can ubiquitously increase GATA6 expression in PDAC cell lines independently of their subtype.

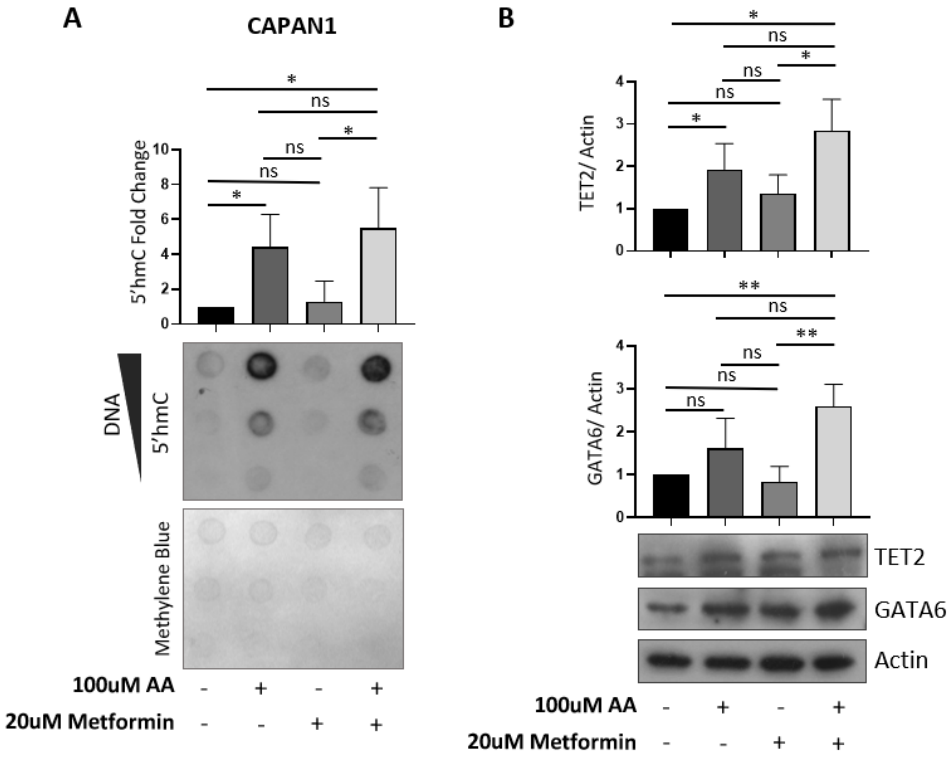


Figure 4.8: **Ascorbic Acid increases GATA6 in CAPAN1 Cells**. **A**) Dot blots from CAPAN1 cells with AA, Metformin, or a combination of the two. AA increased 5'hmC levels but Metformin had no effect **B**) Western blots from treated cells. AA successfully increased GATA6 expression in CAPAN1 cells, while Metformin had no effect.

Finally, to give an indication of the effect of Metformin and AA on healthy pancreatic cells, ductal epithelial cells immortalized with hTERT (DEC-hTERT) were used. Combined treatment could also increase global 5'hmC and slightly increased TET2 levels in these cells, however no changes in GATA6 were seen. This may be due to the fact that GATA6 is already highly expressed in these cell lines and the amount of 5'hmC at the GATA6 promoter has saturated. Therefore, after a certain level additional 5'hmC cannot enhance GATA6 expression further. Because of this, TET2 restoration may selectively target cancer cells due to the fact that they have low levels of 5'hmC.

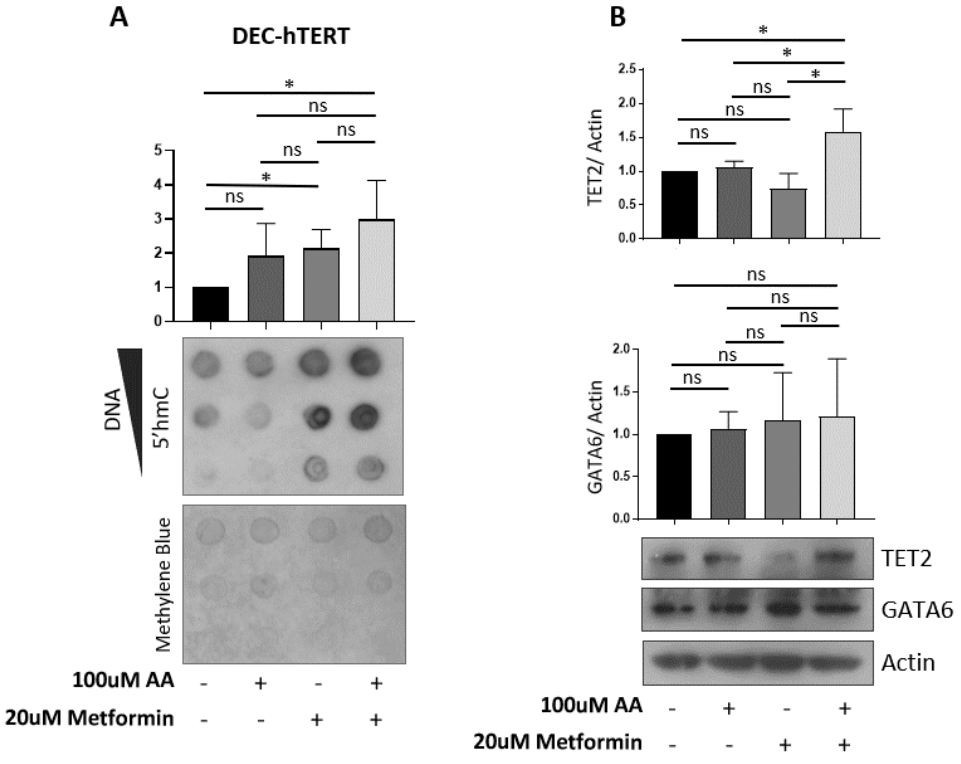


Figure 4.9: **Combined Metformin and Ascorbic Acid can increase TET2, but not GATA6 in Healthy Immortalized Ductal Cells**. **A**) Dot blots from DEC-hTERT cells treated with Ascorbic Acid, Metformin, or a combination of the two. Metformin and AA increased global 5'hmC levels. **B**) Western blots from treated cells. Individual treatment had no effect of TET2 protein levels. Combined treatment resulted in a 50% increase in TET2 levels, but had no effect on GATA6 expression

4.3 Discussion

The previous chapter revealed a correlation between lower 5'hmC and Aggressive molecular subtypes of PDAC; this chapter sought to determine if this relation was causative or purely correlative. To answer this question, TET2 was overexpressed in squamous cell lines, which increased 5'hmC levels and resulted in the re-expression of classical subtype markers such as GATA6. Thus, 5'hmC can actively regulate classical genes, and the decrease in 5'hmC associated with squamous tumours seen in the previous chapter is likely to be a contributing factor towards them becoming more aggressive. These results are particularly interesting given the fact that TET2 stability is regulated by glucose levels via AMPK (Wu et al., 2018). Hyperglycemia is widespread in PDAC patients and patients with diabetes are at a higher risk of developing PDAC, with every 0.56 mM/L increase in fasting blood glucose associated with a 14% increased risk of developing PDAC (Liao et al., 2015). Additionally, many patients without any previous history of diabetes will be hyperglycaemic upon diagnosis of PDAC (Mekan et al., 2006). However, whether this is a consequence of PDAC or is actively driving tumour progression remains unclear. The results presented in this chapter would implicate hyperglycemia in actively promoting PDAC progression by causing TET2 degradation.

The anti-diabetic drug Metformin can increase TET2 stability, and AA can increase TET activity. Individually each of these drugs can increase GATA6 expression in PSN1 cells. Additionally, the combination of both of these drugs together had a synergistic effect and was able to increase GATA6 levels in all three squamous cell lines tested, as well as a classical-pancreatic cell line. This would suggest an increasingly classical subtype is being produced by treating cells, which could potentially make squamous cells less aggressive or prevent classical cells from progressing to squamous. Critically, these results were obtained at physiologically relevant concentrations, making it more likely that these results are clinically translatable.

Both Metformin and AA are known to regulate multiple anti-cancer pathways (Rena et al., 2017)(Vissers and Das, 2018), which raises the question of whether the increases in GATA6 seen are entirely dependent on TET2 or not. While many targets of Metformin and AA have been described, the synergistic effects seen in PSN1 cells would suggest that both drugs are regulating a common target. Additionally, it was shown that regulation of GATA6 by Metformin is AMPK dependent, which narrows down the possibilities for potential targets that could be regulating GATA6 in a TET2-independent manner. One target that fulfills these criteria is mammalian Target Of Rapamycin (mTOR), which is inhibited by Metformin in an AMPK dependent manner (Rena et al., 2017) and activated in response to oxygen (Sengupta et al., 2010), resulting in its inhibition by low dose AA (Yang et al., 2018)(Moretti et al., 2014). Treatment of vascular smooth muscle cells with Rapamycin, which inhibits mTOR, results in phosphorylation of GATA6, leading to GATA6 stabilization and nuclear accumulation, without a change in mRNA levels (Xie et al., 2015). Therefore, increased GATA6 could occur via non-epigenetic mechanisms by Metformin and AA. One way to distinguish between mTOR and TET2 would be to look at GATA6 mRNA as mTOR increases GATA6 stability but not mRNA expression, while TET2 increases GATA6 expression. However, Rapamycin induced differentiation of smooth muscle cells is also dependent of TET2, and mTOR inhibition results in an increase in TET2 at the mRNA and protein level (Liu et al., 2013), suggesting that TET2 is also a mTOR target. Thus, it is difficult to predict to what extent mTOR would regulate GATA6 independently of TET2. To definitively answer this question, a TET2 knockout cell line should be generated, and the effects of Metformin/AA treatment in these cells tested.

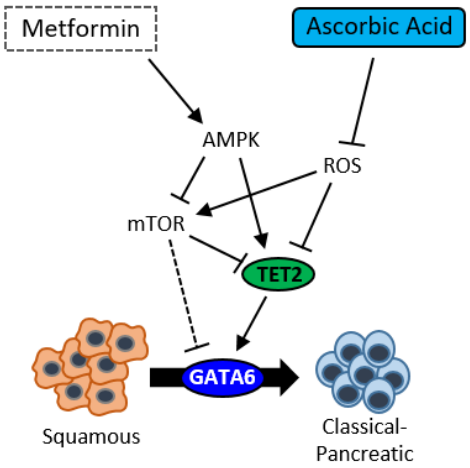


Figure 4.10: **Potential off target effects of Metformin and AA targeting GATA6**. **A**) Schematic showing interaction of Metformin and AA with mTOR, AMPK, TET2 and GATA6.

In conclusion, the results in this chapter implicate TET2 and 5'hmC as a regulator of at least some aspects of molecular subtype identity in pancreatic cancer, and suggest that loss of TET2 can cause classical tumours to become more squamous. Critically, restoration of TET2 in squamous cells makes them more classical, which can be achieved using a combination of metformin and ascorbic acid at physiological levels. Both of these drugs are well tolerated by patients and have been previously investigated in clinical trials for the treatment of PDAC with some success. However, the combination of both drugs has yet to be used in patients and may potentiate their anti-cancer effects.

Chapter 5

Pharmacological Restoration of 5'hmC in vivo

5.1 Introduction

In the previous chapter it was shown that TET2 can actively regulate a panel molecular subtype markers. Furthermore, TET2 can be pharmacologically targeted *in vitro* using a combination of metformin and ascorbic acid at clinically relevant concentrations. However, often *in vitro* studies in cell lines are not translatable to *in vivo* settings. Because of this xenograft studies are commonly used to mimic 3-dimensional tumours by implanting human cell lines into immunodeficient mice (Lai et al., 2017). Typically, cells are implanted subcutaneously. However, while tumours do retain gene expression profiles similar to their site of origin (Chen et al., 2019), PDAC tumours are composed of more than just tumour cells. Indeed, in most cases neoplastic cells are the minority, with the majority of the tumour composed of stromal cells and extracellular matrix (ECM). Therefore, in this chapter the *in vivo* effects of Metformin/AA treatment will be examined using an orthotopic PDAC model.

5.1.1 Aims

This chapter will aim to determine if 5'hmC restoration is possible *in vivo* using Metformin and AA. An orthotopic model of PDAC will be generated using squamous PDAC cell lines in order to recapitulate some of the environmental factors of human tumours and the molecular and phenotypic effects of treating these tumours will be examined.

5.2 Results

5.2.1 Orthotopic Injection of Human PDAC Cell Lines into Nude Mice

In order to determine the effects of Metformin/Ascorbic Acid treatment *in vivo*, an orthotopic PDAC mouse model was used by injecting human PDAC cells directly into the pancreata of nude mice (Fig. 5.1A). This method is more comparable to the human tumour state than conventional xenograft models as it maintains many of the tumour microenvironmental factors important for PDAC progression and treatment (Kim et al., 2009). The approximate location of the pancreas was determined by identifying the silhouette of the spleen, and a small incision was made to expose the pancreas. 500,000 PSN1, PANC1 or MiaPaca2 cells suspended in matrigel were then injected directly into the pancreas. Mice were imaged using magnetic resonance imaging (MRI) 1 week post surgery, and underwent subsequent scans to monitor tumour growth (Fig. 5.1B). PSN1 cells readily grew into large tumours once injected and within two weeks dwarfed the surrounding pancreatic tissue. PANC1 and MiaPaca2 tumours were dramatically smaller by comparison, and at 2 weeks were 1/20th the size of PSN1 tumours ($\sim$5mm^3 vs $\sim$100mm^3 respectively) (Fig. 5.1C). Mice were treated with a combination of Metformin (2.5mg/ml in drinking water) and Ascorbic Acid (2g/kg body weight administered by intraperitoneal [IP] injection) every day for 5 days (Fig. 5.1D). Control mice underwent a mock treatment with a saline IP injection. PSN1 tumour bearing mice began treatment 1 week after surgery, while MiaPaca2 and PANC1 tumours were allowed to grow for longer, and began treatment at 2.5 and 3.5 weeks respectively. Following treatment, tumours were excised and used for histochemical and molecular analysis. Table 5.1 outlines the number of mice for each cell line and treatment group. In one PSN1 tumour bearing mouse, and two MiaPaca2 mice off site tumours could be seen in the abdomen, likely due to leakage of

cells during initial injection. All surgeries, MRI scans and measurements of tumour volume were performed by Simone Lanfredini; a postdoc in the O'Neill lab, with assistance from Sophie Hughes and Asmita Thapa; DPhil students in the O'Neill lab. All subsequent analysis was performed by the author.

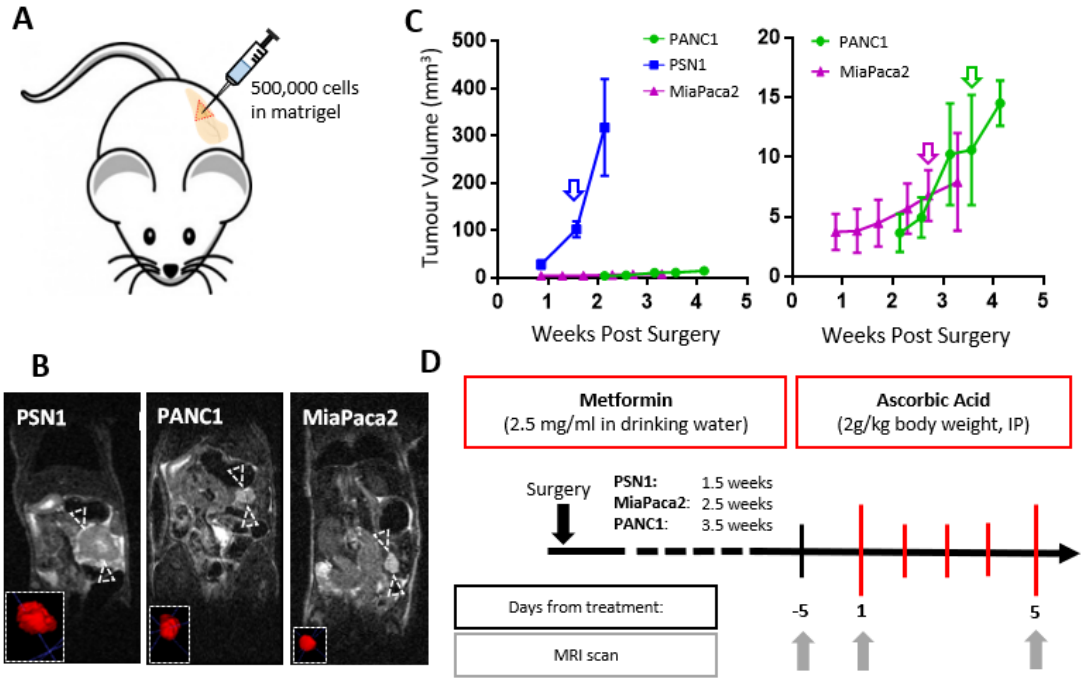


Figure 5.1: **Orthotopic Mouse Model Used for *in vivo* Experiments**. **A**) Schematic showing principle for generation of orthotopic model. PSN1, PANC1 and MiaPaca2 cells were injected directly into the pancreata of nude mice. **B**) Example MRI scans taken at the end of the experiment showing *in situ* tumours in the pancreata of mice. **C**) Left: tumour growth determined by MRI following surgery. Right: the same growth curves with PSN1 tumours excluded. Coloured arrows indicate the start point of treatment. **D**) Dosing scheme for treatment of mice.

Table 5.1: **Orthotopic Mouse Tumours**

Cell Line	No. Treated	No. Mock	Off Site Tumours
PSN1	4	3	1 Treated
PANC1	4	3	None
MiaPaca2	3	3	1 Mock, 1 Treated

Number of treated and control (mock) tumour bearing mice for each cell line.

H&E staining of the pancreata of tumour bearing mice showed PDAC tumours surrounded by largely histologically healthy tissue (Fig. 5.2A). Tumours were composed of cells with large, irregular nuclei interdispersed with smaller cells with oval nuclei (Fig. 5.2B). During PDAC formation, tumour cells release mitogenic substances that stimulate the invasion and proliferation of the surrounding stellate and fibroblast cells, causing them to undergo a desmoplastic reaction to form a dense stroma that is characteristic of PDAC tumours (Rasheed et al., 2012). In return, stromal cells secrete ECM proteins such as collagens and fibronectin, as well as growth factors such as Wnt ligands, that support the growth of the tumour (Maitra and Hruban, 2008)(Thomas and Radhakrishnan, 2019)(Seino et al., 2018). It was hypothesized that the cells with oval nuclei were stroma cells derived from the endogenous mouse tissue. To confirm that this was indeed the case, tumours were stained with an antibody that recognizes the surface of human, but not mouse, mitochondrial membranes. This clearly outlined the cells with large nuclei, but not the small cells (Fig. 5.2C), confirming that these cells are of mouse origin. The largest number of stromal cells was seen in MiaPaca2 tumours, with PANC1 tumours also displaying a large proportion (visual observation, not quantified). PSN1 tumours on the other hand had very few stromal cells, possibly due to the short time frame in which the tumours were grown. To see if these stromal cells are acting in a similar manner to the stroma in human PDAC, tumours were stained with Masson's Trichrome which stains collagen fibres a blue colour. In PANC1 and MiaPaca2 tumours collagen fibres could be seen throughout the tumour covering approximately 10-15% of the tumour section, indicating that a stromal matrix is being formed (Fig. 5.2D). Therefore, these tumours successfully recapitulate some aspects of the PDAC tumour microenvironment. PSN1 tumours on the over hand had very little stroma; consistent with the low number of stromal cells observed.

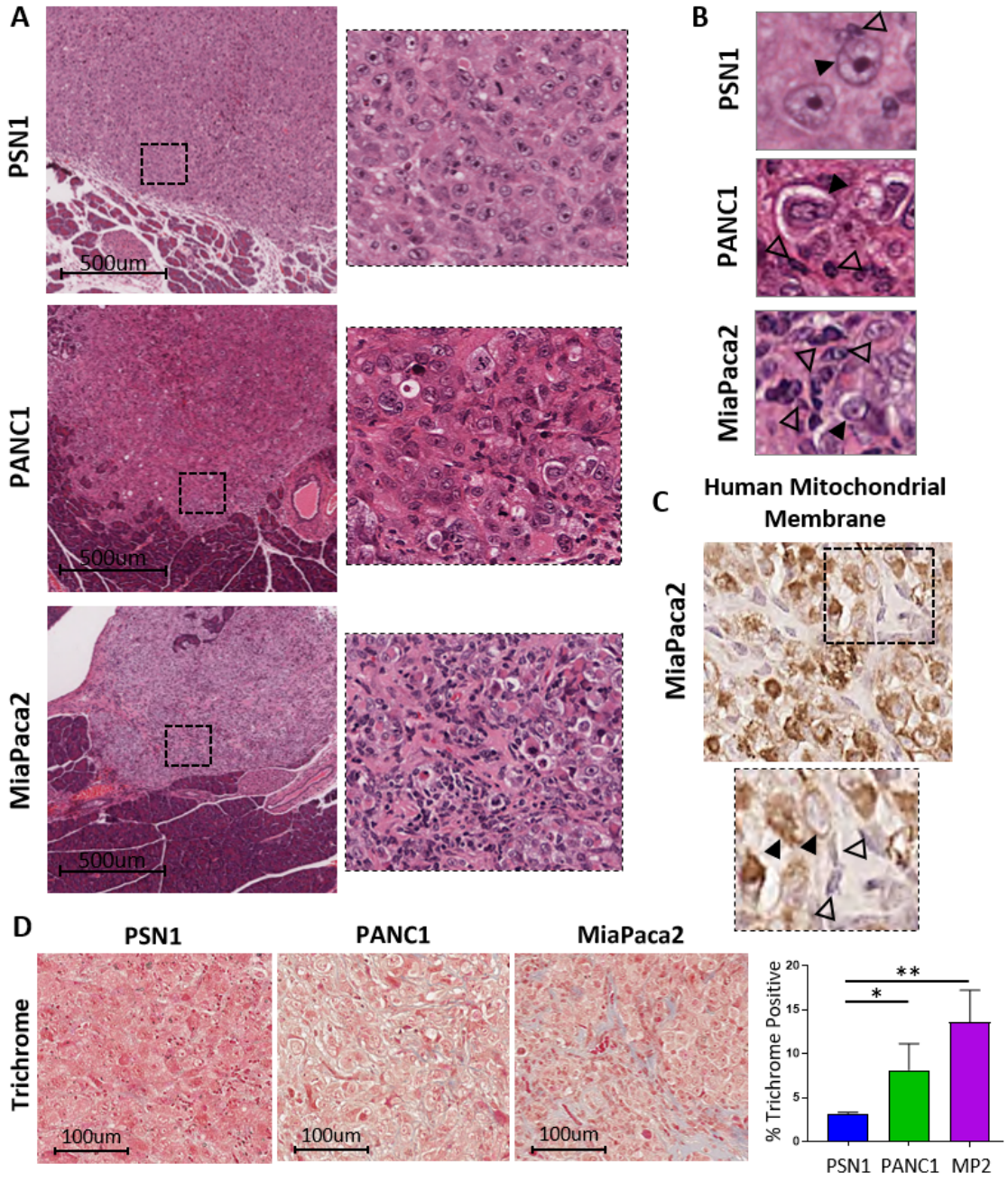


Figure 5.2: **Characterization of Orthotopic Tumours**. **A**) H&E staining of PSN1, PANC1 and MiaPaca2 tumours within mouse pancreata. Right: zoomed regions within tumours. **B**) Higher magnification of tumours showing heterogeneous tumours composed of enlarged, irregular nuclei (black arrowheads) and small, oval nuclei (empty arrowheads). **C**) Representative image of a MiaPaca2 tumour stained with an antibody for human, but not mouse, mitochondrial membrane showing that cells with enlarged nuclei are of human origin while cells with small oval nuclei are not. **D**) Masson's Trichrome staining of tumours. Collagen fibres are stained blue. Collagen can be seen throughout MiaPaca2 and PANC1 tumours, with lower levels found in PSN1 tumours.

5.2.2 Metformin and Ascorbic Acid Increase 5'hmC and GATA6 Levels *in vivo*

In the previous chapter it was shown that combined metformin and AA treatment could increase 5'hmC and GATA6 levels in cell lines. To see if this is also the case *in vivo* global 5'hmC and GATA6 protein levels were determined in orthotopic tumours using IHC. Staining of healthy mouse tissue for 5'hmC showed strong nuclear positivity throughout the pancreas (Fig. 5.3A), similar to results seen in human tissues (see chapter 3). Nuclear staining of the surrounding, morphologically normal, mouse tissue was used as an internal positive control when staining tumours. Tumours cells had much lower levels of 5'hmC than surrounding healthy tissue, however it is unclear whether this difference is due to the difference in species between the mouse pancreas and human cell lines. Nuclear staining was then quantified in mock and treated tumours. Staining was only quantified in tumour cells. Therefore, cells with small oval nuclei were excluded from analysis as they are likely to be mouse derived stromal cells. In all three cell lines tested metformin/AA successfully increased global 5'hmC staining relative to mock tumours (Fig. 5.3B). Next tumours were stained for GATA6. Staining of healthy mouse tissue showed strong nuclear staining in acinar cells, with some cytoplasmic staining (Fig. 5.4A). In untreated tumours nuclear GATA6 staining was weak. Upon metformin/AA treatment GATA6 staining was increased with most cells having moderate staining (Fig. 5.4B). Therefore, combined metformin and ascorbic acid treatment can successfully increase both 5'hmC and GATA6 levels *in vivo*.

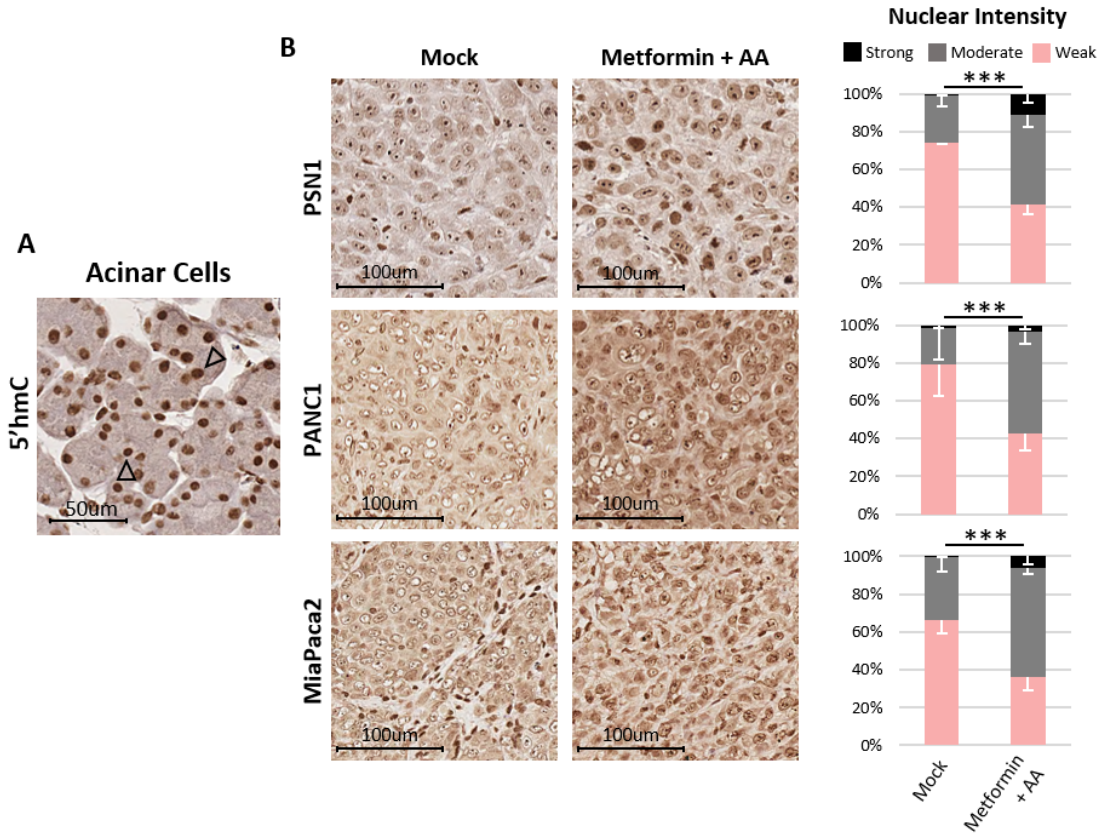


Figure 5.3: **Metformin/AA Increases 5'hmC**. **A**) 5'hmC staining of healthy mouse tissue showing strong nuclear staining (empty arrowheads). **B**) 5'hmC staining of orthotopic tumours for PSN1, PANC1 and MiaPaca2 cells. Quantification of nuclear staining intensity is shown to the right. Stromal cells were excluded from quantification. Metformin/ascorbic acid increased nuclear 5'hmC in all three cell lines.

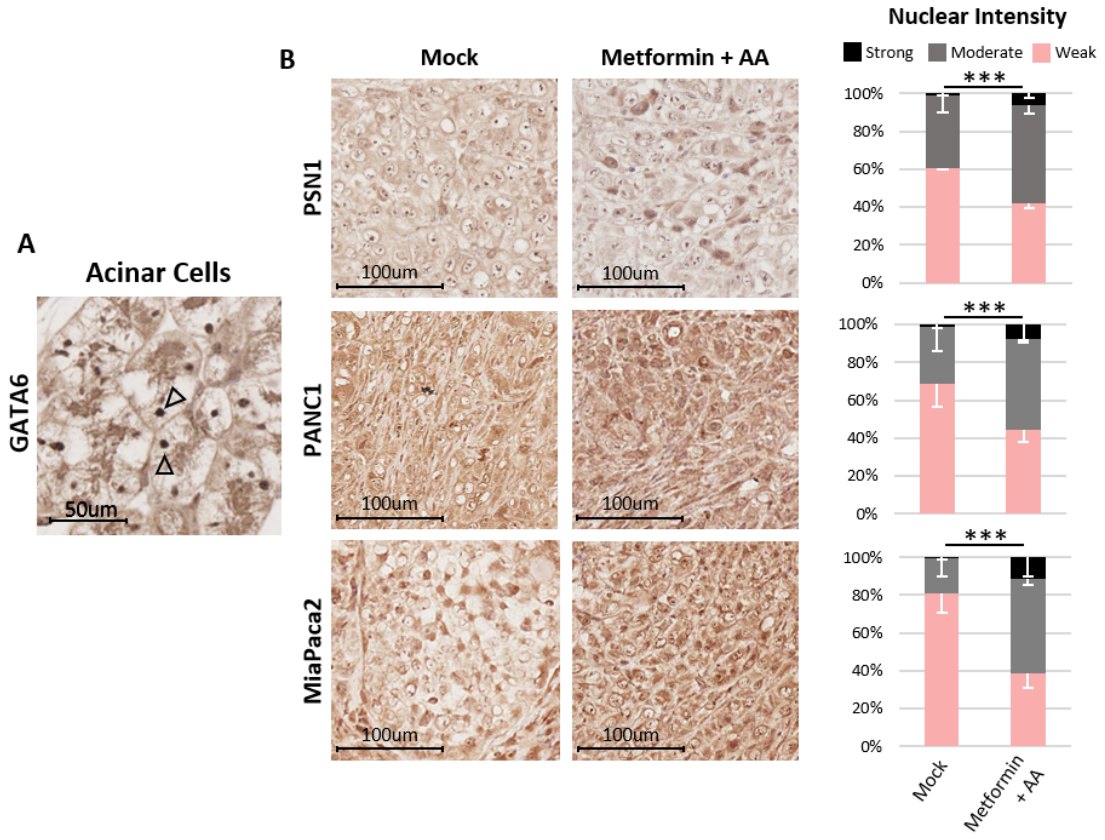


Figure 5.4: **Metformin/AA Increases GATA6**. **A**) GATA6 staining in healthy mouse tissue showing strong nuclear staining in acinar cells (empty arrowheads). **B**) GATA6 staining of orthotopic tumours following metformin and ascorbic acid treatment. Quantification is shown to the right. Stroma cells were excluded from quantification. Metformin/AA increased nuclear GATA6 in all three cell lines.

5.2.3 Restoration of 5'hmC Results in a Partial Switch in Subtype

To better characterize the effects of Metformin and Ascorbic acid treatment 3'RNA-seq was performed on three mock and treated macrodissected PANC1 tumours (Fig. 5.5A). This method involves an enrichment step for the 3'end of RNA transcripts prior to sequencing. Because of this only a single sequencing read is generated for each transcript, directly linking the number of reads for a gene to its expression and reducing the number of sequencing reads required at least 5 fold (Moll et al., 2014). Following 3'RNA-seq, the raw sequencing reads were aligned to the reference genome. The orthotopic model used consists of human tumour cells and mouse stroma cells so the raw reads obtained will contain a mix of mouse and human sequences which, if deconvoluted, can be used to examine tumour and stromal specific changes in gene expression. Raw reads were therefore aligned to both the mouse and human genome. However, the 3'UTR region of mRNAs are highly conserved between species as they are often the sites of miRNA binding (Friedman et al., 2009), which could cause some mouse reads to become improperly aligned to the human genome and *vice versa*. To see if this is be the case, human aligned BAM files were visualized using the Integrated Genomics Viewer (IGV) (Thorvaldsdottir et al., 2013). When examining the GATA6 gene, it was noted that the same missmatches could be seen in a subset of reads. When these were compared to the same region of the mouse genome, the missmatches perfectly corresponded to the mouse sequence, indicating that these reads were incorrectly attributed mouse reads (Fig. 5.5B). To deconvolute these reads, the R programme XenoFilteR was used, which assigns reads present in both the human and mouse datasets to a single species depending on how closely they align to each reference genome (Kluin et al., 2018). The mean number of raw sequencing reads was $\sim$11 million, which resulted in $\sim$2.7 million and $\sim$2.3 million read pairs when aligned to the human and mouse reference genomes respectively. Following filtering, it was

determined that approximately 1 million reads were aligned to the wrong species in each dataset, making up 25-30% of all aligned reads (Fig. 5.5C).

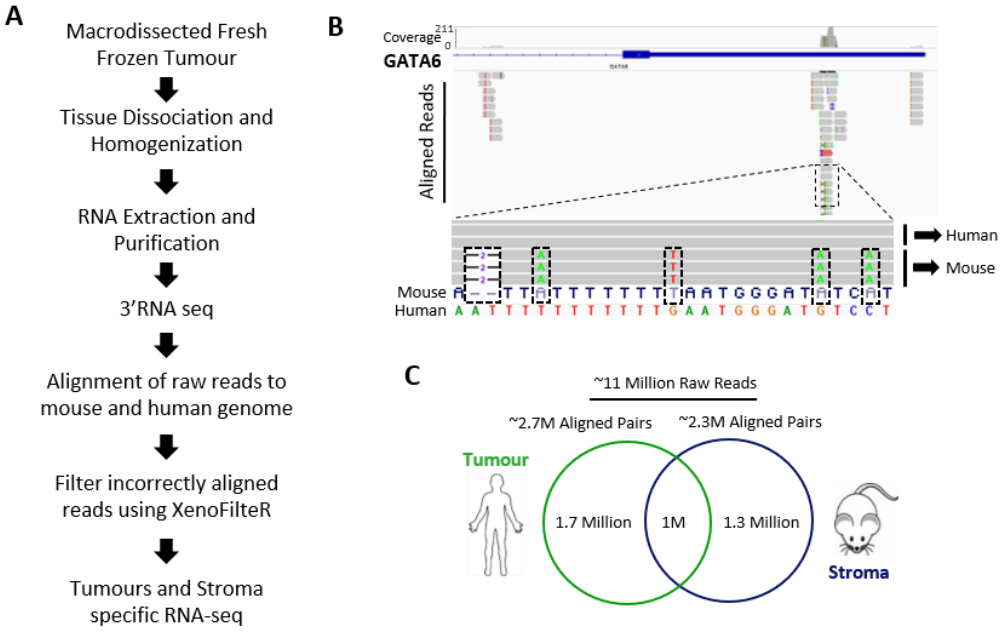


Figure 5.5: **RNA-seq of PANC1 tumours**. **A**) Work flow of RNA-seq experiment. **B**) IGV visualization of RNA-seq reads aligned to the human genome (hg19). The 3' end of the GATA6 gene is shown. Missmatches in some reads perfectly correspond to the mouse sequence for the same region (blue text). **C**) Number of reads attributed to mouse and human genomes following filtering using XenoFilteR.

After successfully deconvoluting tumour and stroma specific signals GATA6 mRNA levels were examined in each compartment. Interestingly, there was an increase in GATA6 mRNA in both mouse and human cells following treatment, indicating that Metformin/AA is active in both the tumour and the stroma (Fig. 5.6A). To see if an increase in GATA6 mRNA and protein within the tumour translates to a switch in the molecular subtype of the cells the relative enrichment of subtype gene sets was determined using GSVA. Treated cells had significantly lower enrichment for the squamous gene set indicating that the tumours are indeed becoming less squamous in identity. However, the remaining gene sets were unchanged suggesting that only a partial switch in molecular subtype has occurred (Fig. 5.6B).

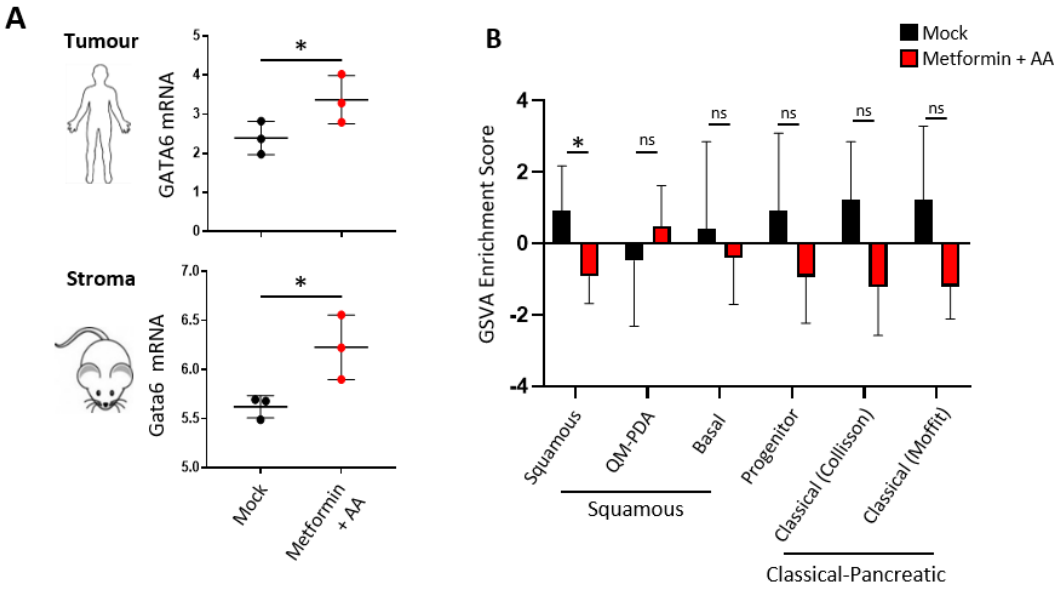


Figure 5.6: **Metformin/AA Induces a Partial Switch in Subtype in PANC1 Cells**. **A**) GATA6 mRNA expression in mouse (stroma) and human (tumour) cells (1 sided t-test). **B**) Enrichment for subtype gene sets for PANC1 tumours. Metformin/AA inhibited the squamous gene set. (1 sided t-test).

5.2.4 Metformin and AA Inhibit Wnt Pathway Signalling

Next, differential gene and gene set analysis was performed between mock and treated tumours (Fig. 5.7). MTOR signalling, a known Metformin target (Rena et al., 2017), was suppressed in treated tumour cells, indicating the treatment was having a function effect as expected. Furthermore, a number of stem cell signatures were downregulated, in line with the idea that Metformin/Ascorbic acid can induce tumour differentiation. Gene sets for signalling pathways linked to PDAC progression were also inhibited in tumour cells, including the VEGF and SRC pathways. Sonic Hedgehog (Shh) signalling on the other hand was activated in response to treatment. Interestingly this may be due to ADP Ribosylation Factor Like GTPase 6 (ARL6) which was one of the most highly expressed genes and is involved in formation of the primary cilia (Liu et al., 2016), which is required for Shh signalling (Goetz et al., 2009).

The Wnt pathway genes WNT3A and R-Spondin 1 (RSPO1) were two of the most

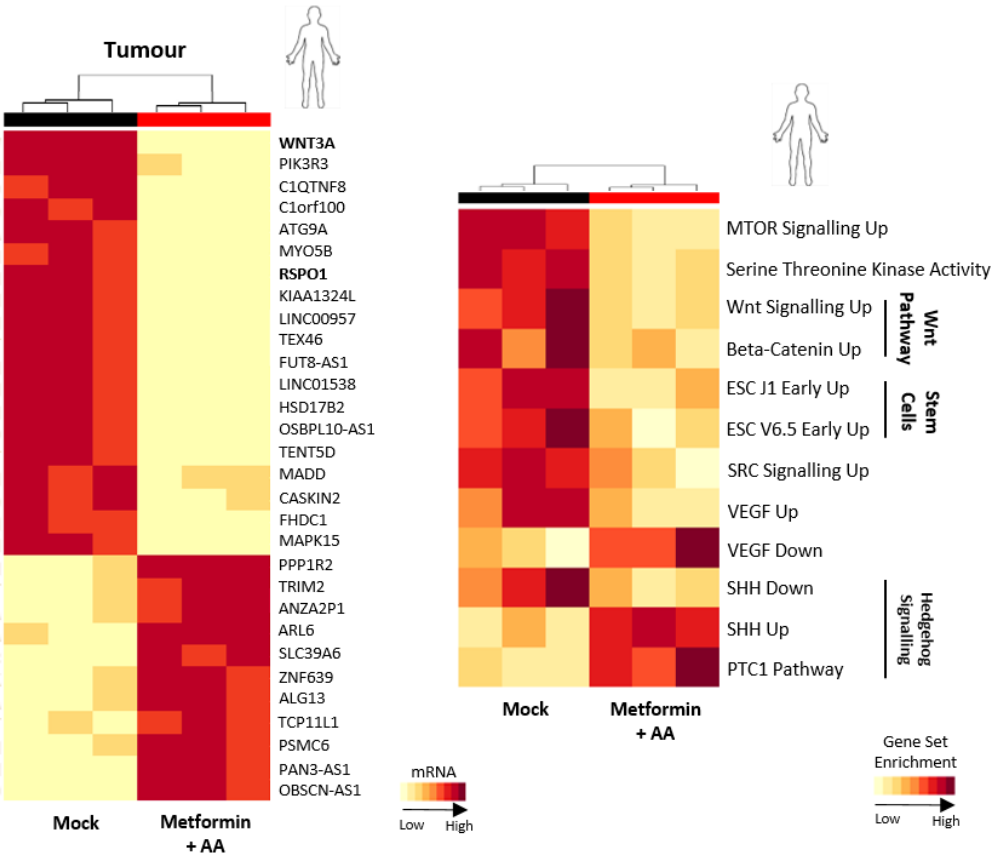


Figure 5.7: **Metformin/ Ascorbic Acid Suppresses Wnt and Stem Cell Signalling Pathways in Tumour Cells**. Left: RNA-seq results from control and treated PANC1 derived tumours. The Wnt pathway genes R-spondin 1 and Wnt3A are downregulated upon treatment. Right: Differentially regulated gene sets between control and treated tumours. Metformin/Ascorbic Acid treatment resulted in suppression of Wnt and Beta-Catenin signalling, as well as multiple stem cell signatures.

significantly downregulated genes upon treatment in tumour cells, and gene sets for the Wnt pathway and Beta-catenin signalling were similarly inhibited (Fig. 5.7A). In PDAC tumours, Wnt ligands are secreted from stromal cells in order to maintain the PDAC stem cell niche. However, tumours that lose GATA6 expression can become independent of these ligands by producing their own Wnts, including Wnt3A, Wnt7B and Wnt10A (Seino et al., 2018). This has been linked to increased tumour aggressiveness and a more squamous subtype as tumour cells are no longer reliant on the stroma to support them (Arensman et al., 2014)(Seino et al., 2018). Therefore, the decrease

in Wnt signalling and stem cell signatures may be a downstream effect of GATA6 suppressing autonomous Wnt production in the tumours. Indeed, in the Oxford cohort there was a highly significant negative correlation for both GATA6 and TET2 mRNA with Wnt pathway genes, including RSPO1, Wnt3A, Wnt7B and Wnt10A; consistent with the idea that GATA6 and TET2 are inhibiting Wnt Signalling in PDAC tumours (Fig. 5.8A). Interestingly, when differential gene expression analysis was performed in stromal cells, Wnt7b was also found to be significantly downregulated, suggesting that Metformin/AA can inhibit both autocrine and paracrine Wnt signalling (Fig. 5.8B).

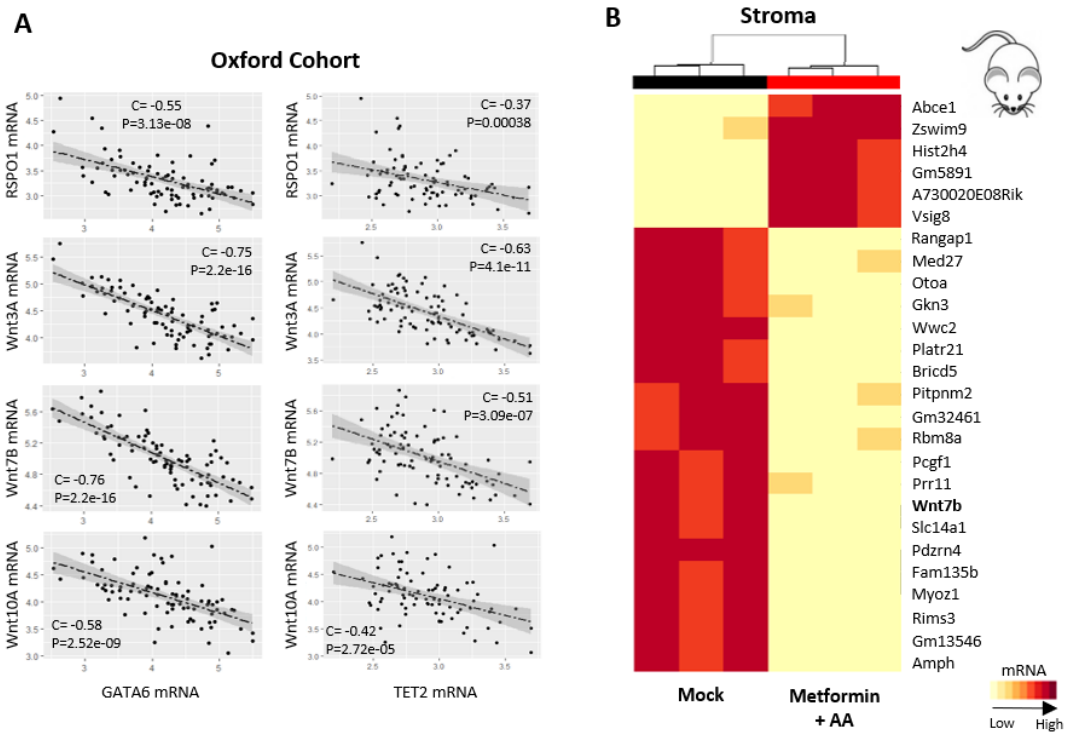


Figure 5.8: **Metformin/ Ascorbic Acid Suppresses Wnt7b in Stroma Cells**. **A**) Differential gene expression in mouse cells between mock and treated tumours **B**) Correlation between (left) GATA6 and (right) TET2 mRNA with Wnt pathway genes in the Oxford Cohort. P-value and Pearson's correlation are shown in the bottom left or top right of each panel.

To confirm that Wnt signalling was being inhibited by Metformin/AA, tumour sections were stained for β-catenin. As a positive control to test correct antibody

staining, a section of healthy mouse intestine was stained, as the role of the Wnt pathway during intestinal stem cell maintenance and differentiation has been highly characterized (Reviewed by Kretzschmar and Clevers (2017)). Staining of intestinal crypts showed nuclear staining in intestinal stem cells (columnar cells at the base of crypts), but membranous staining in all other differentiated cell types; consistent with the established localization of β-catenin in these cell types (Fig. 5.9A). In healthy pancreatic tissue, β-catenin was localized at the membranes of both acinar and ductal cells (Fig. 5.9A). In contrast, nuclear β-catenin could be seen in the nuclei of mock tumour cells, with 40% of PSN1 and 90% of PANC1 and MiaPaca2 tumour cells having moderate or strong nuclear staining (Fig. 5.9B). Metformin/Ascorbic Acid treatment resulted in a 50% decrease in the proportion on PSN1 and PANC1 cells with moderate/strong nuclear staining, while a smaller decrease of 20% was seen in MiaPaca2 cells. Thus, Metformin/Ascorbic acid treatment can indeed inhibit downstream Wnt pathway signalling.

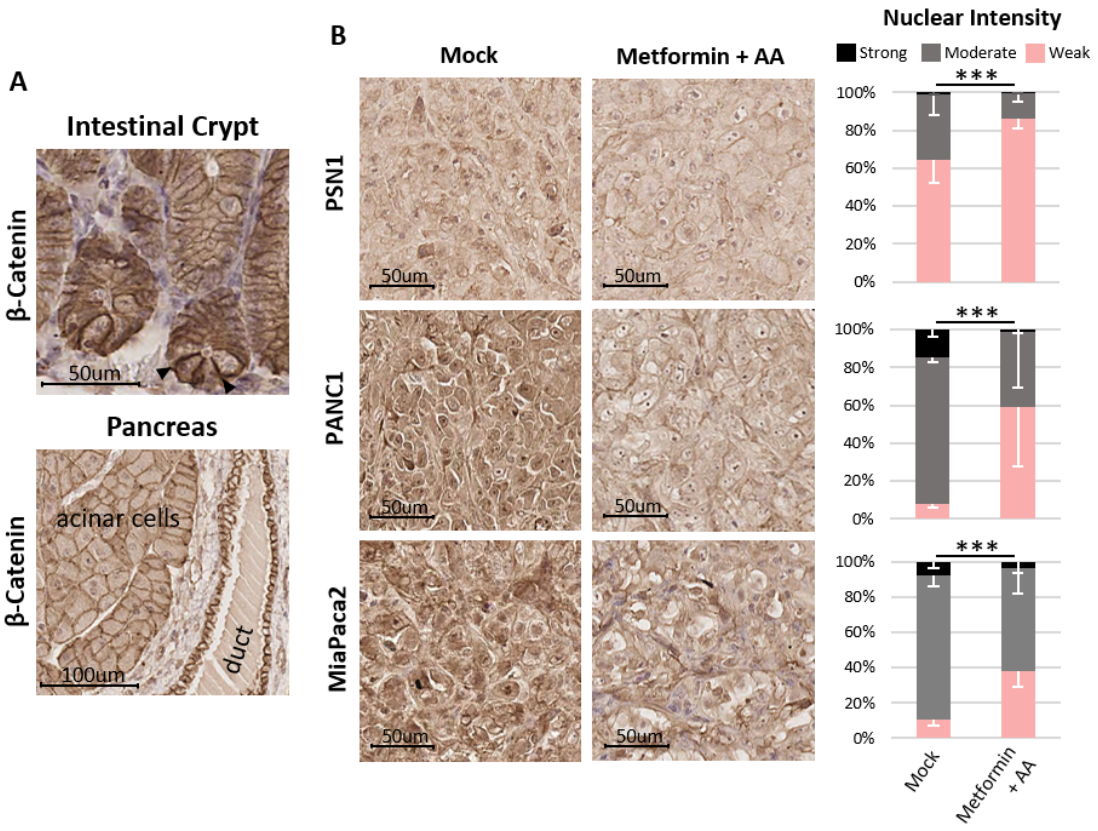


Figure 5.9: **Metformin/ Ascorbic Acid Inhibits nuclear β-Catenin in Tumour Cells**. **A**) Positive controls for Beta-catenin staining showing (top) intestinal crypts and (bottom) healthy pancreatic tissue. β-catenin is nuclear in intestinal stem cells (black arrowheads) but membranous in differentiated intestinal and pancreatic cells. **B**) β-Catenin staining in Orthotopic tumours. Nuclear staining can be seen in mock tumours, which is reduced when tumours are treated with Metformin and Ascorbic acid. Right: Quantification of nuclear staining in tumours.

5.2.5 Metformin/AA Inhibit Tumour Proliferation

Finally, the phenotypic consequences of Metformin/AA treatment were examined. While loss of GATA6 marks highly aggressive tumours, GATA6 itself has been shown to have oncogenic effects in PDAC and can drive cell proliferation in cell lines (Kwei et al., 2008). Conversely, inhibition of downstream Wnt signalling can suppress the proliferation of both Wnt dependent and Wnt autonomous organoids (Seino et al., 2018). Therefore it is difficult to predict the effects of GATA6 restoration in squamous tumours. Tumour bearing mice underwent MRI scans on the first and last days of treatment, allowing for the change in tumour volume during this time to be measured. Mock treated PSN1 tumours doubled in size over the 5 days of treatment. In Metformin/Ascorbic acid PSN1 treated tumours on the other hand there was only a 50% increase in tumour volume. However, no difference was seen in PANC1 and MiaPaca2 tumours (Fig. 5.10A).

As PANC1 and MiaPaca2 tumours were much smaller than PSN1 tumours any changes in tumour volume will be more difficult to pick up on MRI. Furthermore, as tumours margins must be manually outlined on MRI scans there is likely to be a high margin of error in any volume measurements, which will disproportionately effect smaller tumours. In order to better assess tumour proliferation, tumour sections were stained for ki67, which is active in all cycling cells, but absent in quiescent cells (Bruno and Darzynkiewicz, 1992). The proportion of cycling tumour cells was higher in PSN1 tumours compared to PANC1 and MiaPaca2 ($\sim$60% vs $\sim$50%); consistent with the higher growth rate of these tumours. In all treated tumours there was a 20-40% decrease in the proportion of ki67 positive tumour cells, indicating the Metformin/Ascorbic acid can inhibit tumour proliferation in all three cell lines *in vivo* (Fig. 5.10B).

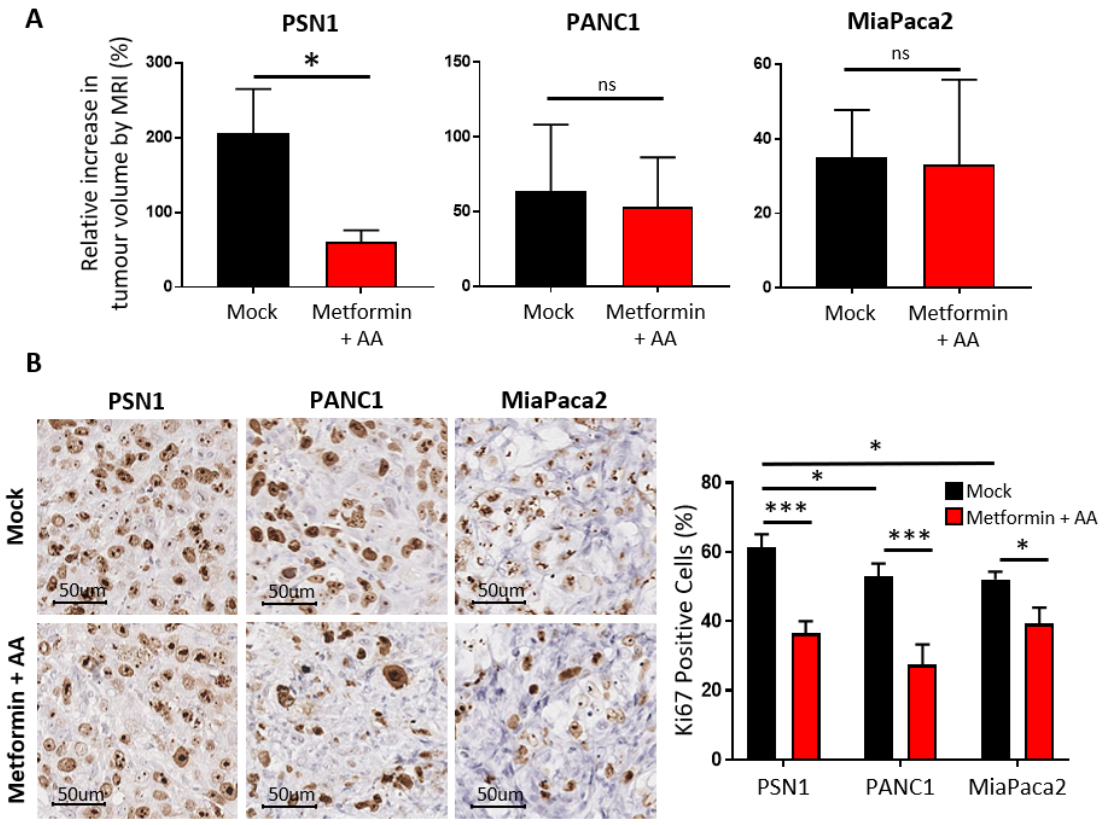


Figure 5.10: **Metformin and Ascorbic Acid Inhibit Tumour Proliferation**. **A**) Relative increase in tumour volume during the time of treatment in mock and treated mice. Mice underwent MRI scans before and after treatment. **B**) ki67 staining of tumours. Metformin/AA decreased the proportion of ki67 positive cells in the tumour.

5.3 Discussion

In this chapter an orthotopic mouse model of PDAC was used in order to determine if the results seen in the previous chapter are applicable to *in vivo* tumours. Injection of PDAC cell lines into the pancreata of mice resulted in the generation of 3-dimensional tumours composed of a mix of human tumour cells and mouse stroma cells. Additionally, collagen fibres could be seen through the tumours, indicating that an ECM was being formed by the stroma cells. Thus, these tumours successfully recapitulate aspects of the tumour microenvironment seen in human PDAC tumours. It should be noted that tumours were generated in nude mice. Therefore, while the immune system is increasingly being recognized as a regulator of tumour development and progression (Inman et al., 2014), it is not represented within this model. Interestingly, combined metformin/AA treatment increased global 5'hmC and nuclear GATA6 protein levels in this model, suggesting that this combination may be able to increase GATA6 in patient tumours in order to make them less squamous.

RNA-seq was carried out on tumours to better characterize the effects of treatment. This showed that GATA6 mRNA was increasing in both the tumour and stroma cells. However, this was only sufficient to inhibit squamous gene expression and not to reinstate a classical-pancreatic gene signature. Therefore only a partial switch in molecular subtype occurred. In the previous chapter it was stated that GATA6 increase by metformin and AA could be due to TET2, which regulates GATA6 expression, or mTOR, which regulates GATA6 stability. The fact that GATA6 mRNA is increased suggests that at least some of the GATA6 increase seen in treated tumours is due to restoration of TET2. However, gene sets for mTOR signalling were also inhibited following treatment. Therefore, it is likely that metformin/AA can increase GATA6 levels in PDAC tumours through multiple mechanisms.

The RNA-seq analysis showed that Wnt ligand expression was being downregulated. Additionally, IHC for β-catenin showed that downstream Wnt signalling was

also suppressed. Interestingly, Seino et al. (2018) showed that Wnt niche requirements were variable between tumours, with some tumours being dependent on the stroma to supply them with Wnt ligands and others capable of activating their own Wnt signalling. Metformin/AA suppressed Wnt expression in both stroma and tumour cells, suggesting that both paracrine and autocrine signalling are being suppressed. Given that the ability of organoids to become Wnt autonomous is due to GATA6 loss, the inhibition of Wnt signalling seen in tumours may be due to GATA6 restoration. Seino et al. (2018) also found that Wnt niche independency was likely to be epigenetically acquired as it took 2-3 weeks before KRAS, TP53, SMAD4 and CDKN2A mutant organoids could grow in the absence of supplemented Wnt. Taken together, the results here and in the previous two chapters suggest that TET2 is the epigenetic mechanism that acts upstream of GATA6 loss, Wnt independency and acquisition of a squamous subtype.

Chapter 6

SMAD4 Regulates 5'hmC by Promoting TET2 Expression

6.1 Introduction

The previous chapters implicated TET2 as a regulator of PDAC molecular subtype and suggested by extension that any factor that inhibits TET2, such as diabetes and hyperglycemia, can contribute to progression to a squamous subtype. The integrated molecular analysis in results chapter 1 highlighted a link between a low 5'hmC epigenetic state and SMAD4 mutations in both the Oxford and PAAD TCGA cohorts. In this final chapter the association between SMAD4 and 5'hmC will be explored further.

6.1.1 SMAD4 Loss in PDAC

SMAD4 was originally identified as a tumour suppressor due to its frequent deletion in pancreatic cancer, resulting in its alternative name of Deleted in Pancreatic Cancer 4 (DPC4) (Hahn et al., 1996). SMAD4 mutations are found in 19-50% of PDAC tumours (Blackford et al., 2009) and are the third most frequent PDAC driver mu-

tation after KRAS and TP53 (Aguirre et al., 2018). Because of this SMAD4 is one of the most extensively studied prognostic markers for PDAC, with a meta-analysis involving over 1700 patients finding SMAD4 mutation to be significantly associated with poor patient outcome (Shugang et al., 2016).

Functionally, SMAD4 is a member of the TGF-β signalling pathway (reviewed by Ahmed et al. (2017) and Zhao et al. (2018). Activation of the TGF-β receptor results in phosphorylation of receptor-regulated Smad proteins (R-Smads); Smad2 and Smad3. Phospo-Smad2/3 then interacts with Smad4 to form a heterotetromeric complex that translocates to the nucleus to drive the expression of a number of tumour suppressive genes such as E-Cadherin and Death Associated Protein Kinase (DAPK). Interestingly, phospho-SMAD2/3 can activate some target genes following TGF-β activation in the absence of SMAD4. In PDAC cells SMAD2/3 can promote the expression of Snail in a SMAD4-dependent manner and SRY-Box 4 (Sox4) in a SMAD4-independent manner. Snail then represses the gastrointestinal lineage master regulator Kruppel Like Factor 5 (KLF). In the absence of Sox4 KLF5 promotes apoptosis. However, if KLF5 is present the two cooperate to drive a progenitor-like identity (David et al., 2016). Thus, when SMAD4 is present TGF-β can have tumour suppressive effects, while SMAD4 mutations cause TGF-β to become oncogenic. Furthermore, TGF-β receptors can active additional signalling pathways in an oncogenic, Smad-independent manner. For example, TGF-β1 can activate Ras, resulting a Erk pathway activation and leading to EMT (Yan et al., 1994)(Mulder and Morris, 1992). It is perhaps because of this that TGF-β signalling frequently has a dual role in tumour progression; initially acting as a tumour suppressor and later as an oncogene once SMAD4 is lost (Lebrun, 2012).

6.1.2 Aims

In chapter 3 low 5'hmC levels were associated with SMAD4 mutations. In this chapter this association will be explored further by examining the effects of SMAD4 modulation on TET2 and 5'hmC levels.

6.2 Results

6.2.1 SMAD4 Promotes TET2 Transcription

In the integrated analysis performed in Chapter 3, Figure 8, the aggressive Cluster 1 characterized by low 5'hmC levels was also associated with an increased incidence of SMAD4 mutations, along with mutations in TP53, RB1 and KRD. The association between SMAD4 and 5'hmC was strengthened further in Chapter 3, Figure 11 when SMAD4 mutations were again found to be associated with a low 5'hmC methylome in the PAAD TCGA cohort. Suggesting that SMAD4 loss may contribute to loss of 5'hmC. In line with this hypothesis, patients with SMAD4 mutations had lower levels of 5'hmC at tumour specific probes that patients with wild type SMAD4 (Fig. 6.1A). Given that SMAD4 functions as a transcription factor the association between SMAD4 and TET2 mRNA was examined, revealing a positive correlation between the two in both the Oxford and PAAD TCGA cohort (Fig. 6.1B) and suggesting that TET2 may be a SMAD4 target gene.

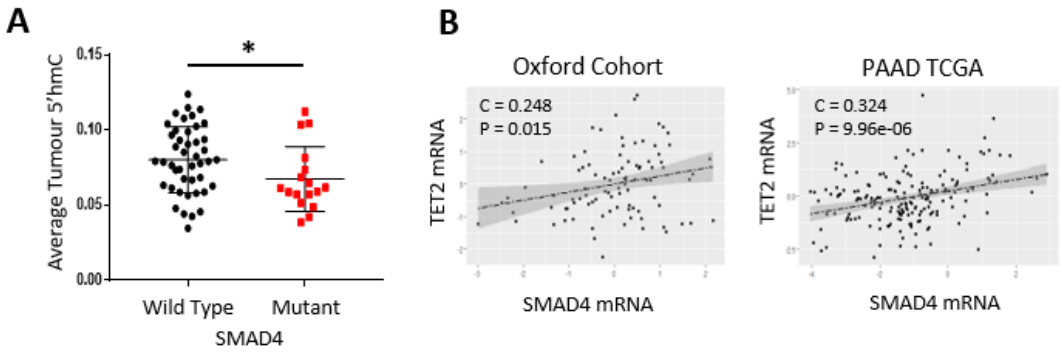


Figure 6.1: **SMAD4 Correlates with TET2**. **A**) 5'hmC levels in SMAD4 WT and mutant tumours. **B**) Correlation between SMAD4 and TET2 mRNA in (left) Oxford and (right) PAAD TCGA cohorts.

In order to test if SMAD4 can indeed regulate TET2, FLAG-SMAD4 was overexpressed in PSN1 cells, which have a homozygous SMAD4 deletion (Moore et al., 2001)(Deer et al., 2010). SMAD4 overexpression increased TET2 protein levels (Fig. 6.2A). Additionally, this upregulation of TET2 was functionally relevant as global 5'hmC levels were also increased (Fig. 6.2B).

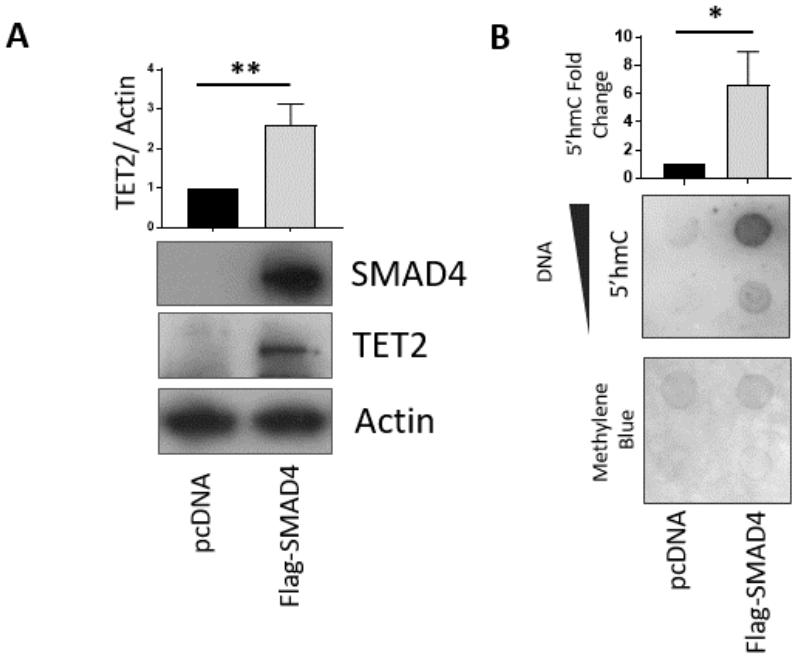


Figure 6.2: **SMAD4 Promotes TET2**. **A**) FLAG-SMAD4 was overexpressed in SMAD4-null PSN1 cells. SMAD4 overexpression increased SMAD4 protein levels as well as (**B**) global 5'hmC.

Next, the effects of SMAD4 loss were examined. SMAD4 was knocked down by siRNA in DEC-hTERT and MiaPaca2 cells. DEC-hTERT are immortalized ductal epithelial cells and therefore should not harbour any driver mutations, while MiaPaca2 cells are SMAD4 WT. In both cell lines SMAD4 knockdown decreased TET2 at both the mRNA and protein level (Fig. 6.3A,B), indicating that SMAD4 can promote TET2 transcription. Thus, mutation of SMAD4 in PDAC tumours may contribute to 5'hmC loss.

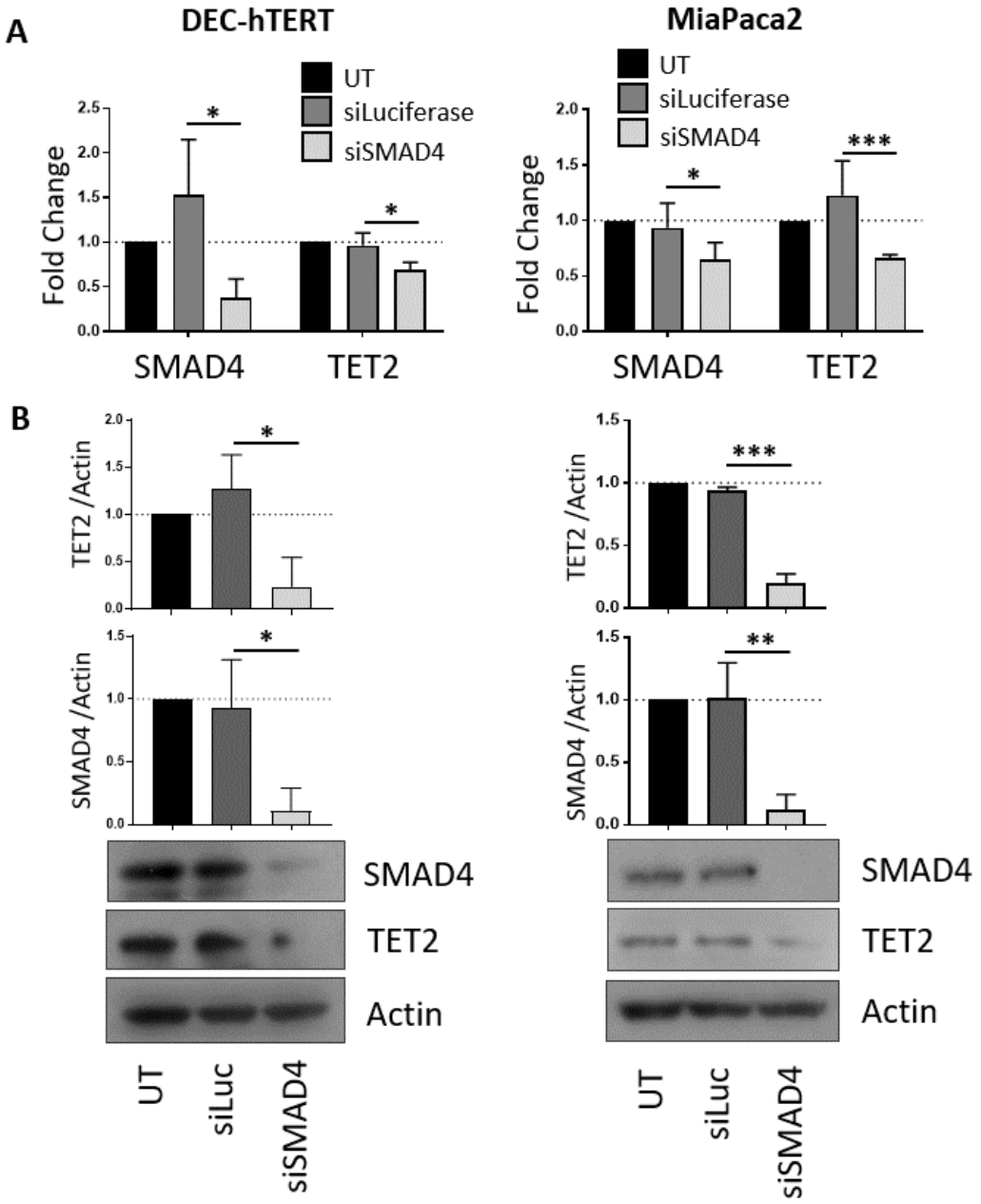


Figure 6.3: **SMAD4 Knockdown Decreases TET2 and 5'hmC**. Knockdown of SMAD4 in (left) DEC-hTERT and (right) MiaPaca2 cells. TET2 levels were determined by (**A**) qPCR and (**B**) western blot.

SMAD4 may promote TET2 transcription by either directly activating the TET2 promoter, or by influencing some other factor that directly activates TET2. To determine if SMAD4 can bind to the TET2 gene a publicly available SMAD4 CHIP-seq dataset performed in an ovarian cancer cell line by Kennedy et al. (2011) was examined. This revealed a SMAD4 binding peak at the distal end of the TET2 gene when cells were stimulated with TGF-β (Fig. 6.4), confirming that SMAD4 can bind directly to the TET2 gene region, though not at the promoter. The region around this binding peak was explored further using datasets from the Encyclopedia of DNA Elements (ENCODE)(Davis et al., 2018), which contains a plethora of CHIP-seq datasets for various transcription factors and histone marks in multiple cell lines (the datasets used combine data from multiple cell lines together). Interestingly, the distal end of the TET2 gene close to where the SMAD4 binding peak was found contains a region with multiple DNase hypersensitivity sites, a marker of accessible chromatin, and was dense with bound transcription factors, suggesting that this region may be an enhancer. This idea is supported by the presence of H3K4 monomethylation (H3K4me1), which marks all enhancers (Heintzman et al., 2007), and H3K27 acetylation (H3K27Ac) which is present at active enhancers (Creyghton et al., 2010) (Fig. 6.4). Given that SMAD4 can promote TET2 transcription, it is possible that this enhancer interacts with the TET2 promoter to bring SMAD4 into contact with the promoter to drive TET2 transcription. However, this enhancer could also interact with a different gene so it is not possible to determine what its target is for sure with the information available. Therefore, SMAD4 may directly promote TET2 transcription by binding to an enhancer that stimulates TET2, but may also promote TET2 transcription by some other mechanism.

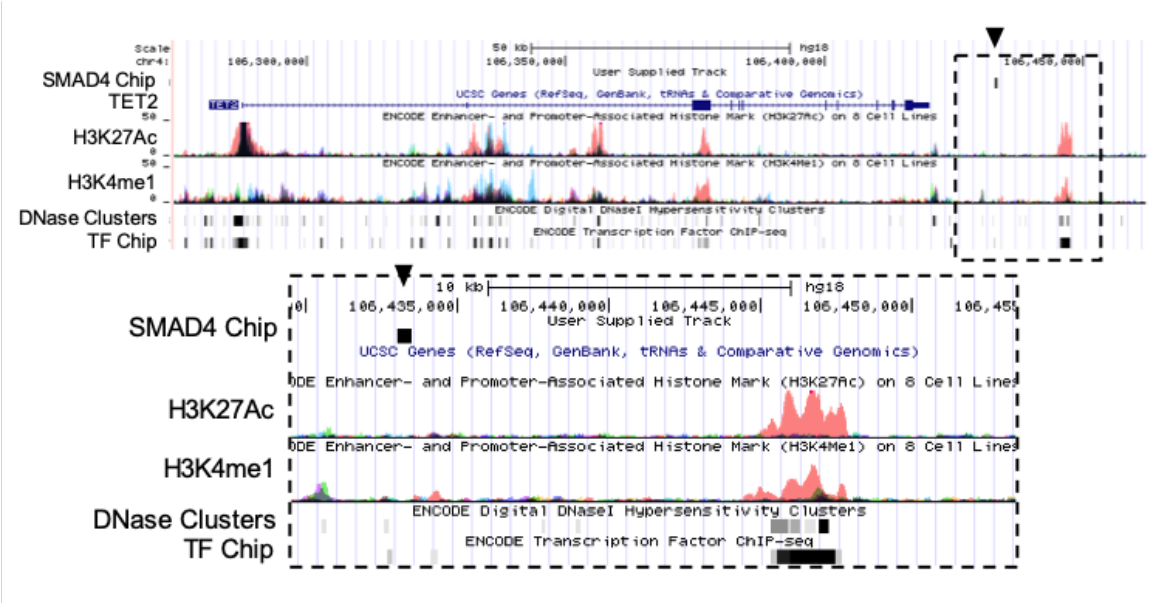


Figure 6.4: **SMAD4 Binds to an Enhancer at the TET2 Gene**. SMAD4 CHIP-seq binding peaks from a dataset generated by Kennedy et al. (2011). A single SMAD4 peak can be seen at the distal end of the TET2 gene (black arrowhead). ENCODE datasets for H3K27Ac, H3K4me1 as well as combined transcription factor (TF) CHIP-seqs are also shown, as well as DNase clusters. Combined H3K27Ac and H3K4me1 mark active enhancers. SMAD4 binds close to an enhancer region at the TET2 gene, but it is unknown if this enhancer regulates the TET2 promoter.

6.2.2 SMAD4 Regulates Molecular Subtype Markers

If SMAD4 is partially responsible for maintaining TET2 levels, then altering SMAD4 levels should influence molecular subtype markers such as GATA6 and PDX1 in a similar manner to TET2. This is not the first time that SMAD4 has been linked to PDAC molecular subtypes; Moffitt et al. (2015) initially described classical tumours as having significantly higher SMAD4 mRNA than basal. A similar pattern was also seen in the Oxford cohort, with squamous tumours having significantly lower SMAD4 mRNA that classical-pancreatic (Fig. 6.5A), as well as significant positive correlations between SMAD4 and GATA6 mRNA (Fig. 6.5B). SMAD4 mutations were also slightly enriched in squamous tumours, but not significantly so ($P = 0.09$) (Fig. 6.5C).

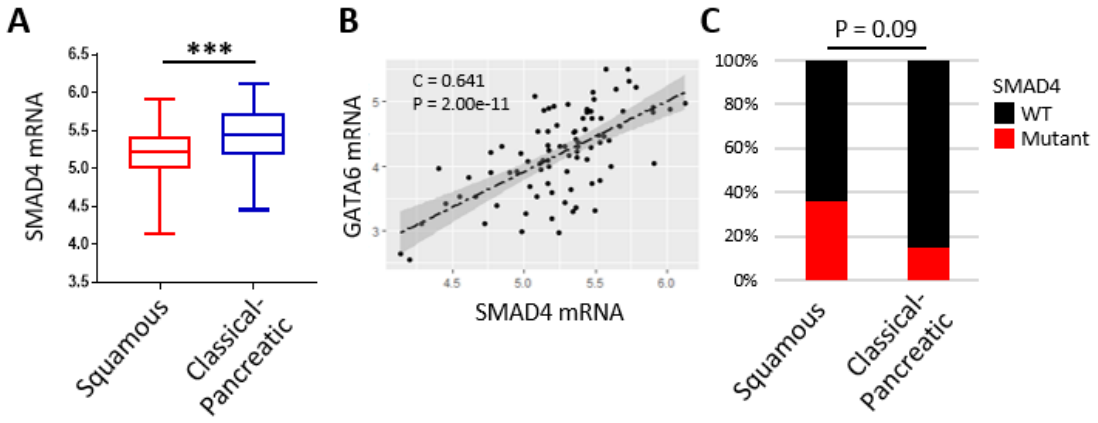


Figure 6.5: **SMAD4 Correlates with Molecular Subtype**. **A**) SMAD4 mRNA in squamous and classical-pancreatic tumours. **B**) Correlation between SMAD4 and GATA6 mRNA in the Oxford Cohort. **C**) Distribution of SMAD4 mutations in squamous and classical-pancreatic tumours. SMAD4 mutations are slightly enriched in squamous tumours but not significantly so.

Interestingly, SMAD4 overexpression increased GATA6 and PDX1 protein levels in PSN1 cells (Fig. 6.6A), indicating that SMAD4 can actively regulate subtype genes in a similar manner to TET2. CAPAN1 cells have a homozygous deletion in SMAD4 so the effects of SMAD4 loss in a classical-pancreatic cell line could not be tested. However, SMAD4 knockdown in DEC-hTERT cells resulted in lower GATA6 and PDX1 expression (Fig. 6.6B), indicating that loss of SMAD4 can cause loss of endodermal differentiation marker expression. These results suggest that loss of SMAD4 in PDAC tumours can promote a more squamous subtype, possibly due to loss of TET2.

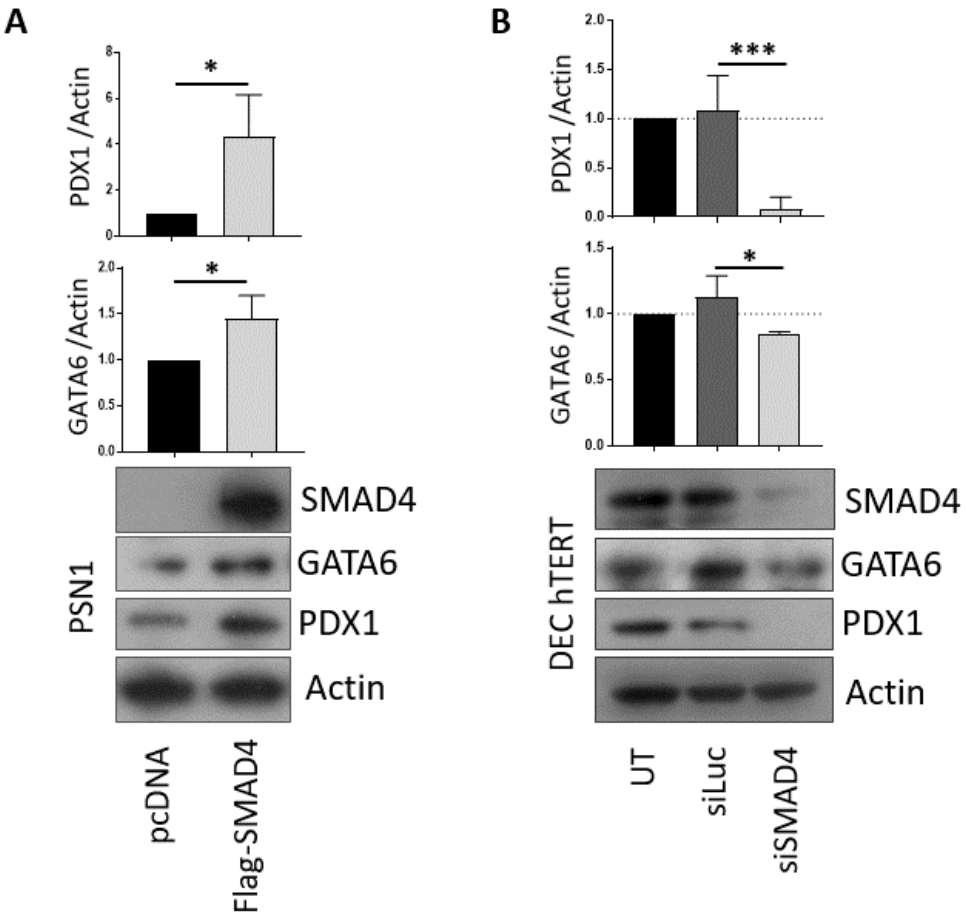


Figure 6.6: **SMAD4 Regulates Molecular Subtype Markers**. **A**) Expression of GATA6 and PDX1 following TET2 overexpression in PSN1 cells and (**B**) TET2 knockdown in DEC-hTERT cells. Note that these western blots are the same as the ones from Figure 6.2A and 6.3A and so the SMAD4 and Actin panels are the same as those figures.

Finally, the transcriptional effects of SMAD4 on molecular subtype identity was analyzed using a publicly available mRNA dataset by David et al. (2016). David et al isolated pancreatic tumour cell lines from mice with mutant KRAS, SMAD4 and CDKN2A (also known as INK4A), referred to as KSIC (KRAS, SMAD4, INK4A, Cre) mice. SMAD4 was then reintroduced into cells using a doxycycline inducible construct (Tet-on SMAD4) (Fig. 6.7A). The relative enrichment of molecular subtype gene sets was determined in KSIC cell lines and KSIC + SMAD4 following 24 hours of SMAD4 induction using GSVA. Only two replicates are available, therefore it is assumed that results have a normal distribution when performing statistical analysis (two-sided T-Test). SMAD4 restoration reduced squamous and QM-PDA gene set enrichment in both replicates indicating that SMAD4 can induce transcriptional changes consistent with a less squamous subtype (Fig. 6.5B). However enrichment of the remaining gene sets varied between replicates, indicating that, similarly to metformin and ascorbic acid treatment, SMAD4 restoration alone is not enough to cause a complete subtype switch. Thus, the results in this chapter identify a novel tumours suppressive function of SMAD4 in maintaining 5'hmC levels and maintaining a classical subtype identity.

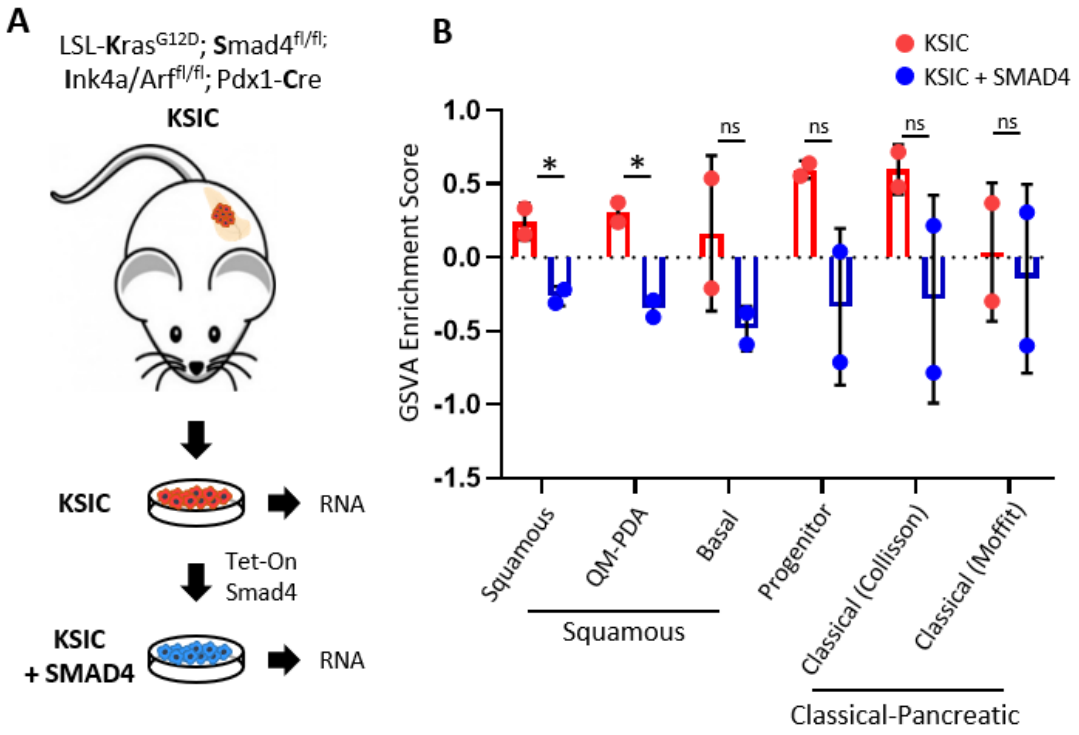


Figure 6.7: **SMAD4 Regulates Molecular Subtype Gene Sets**.**A**) Schematic showing experiments carried out by David et al. (2016). Cell lines were isolated from mice with mutant KRAS, SMAD4 and INK4A (KSIC). SMAD4 expression was restored using an inducible construct (KSIC + SMAD4). **B**) Relative enrichment for subtype gene sets in KSIC and KSIC + SMAD4 cell lines using GSVA. SMAD4 expression inhibited squamous and QM-PDA gene sets but did not alter classical-pancreatic gene sets. Only two replicates are available so it is assumed that results have a normal distribution for statistical analysis.

6.3 Discussion

SMAD4 is one of the most extensively studied genes in PDAC due to its frequent mutation and links to tumour aggressiveness. Typically, loss of SMAD4 is thought to promote oncogenesis through loss of TGF-β associated genes. However, the results in this chapter implicate an additional role of SMAD4 in maintaining the epigenetic state of pancreatic cells. SMAD4 can regulate both the mRNA and protein levels of TET2, indicating that SMAD4 can promote TET2 transcription. This would suggest that loss of SMAD4 in PDAC tumours would contribute to the epigenetic silencing of tumour suppressor genes due to loss of proper 5'hmC regulation. Strikingly, both TP53 and KRAS have also been linked to 5'hmC regulation, though not in PDAC. In mouse ESCs (mESCs) TP53 promotes TET1 and TET2 expression, and TP53 mutation results in loss of TETs leading to widespread hypermethylation and spontaneous differentiation (Tovy et al., 2017). Notably TP53 mutations are also enriched in squamous tumours (Bailey et al., 2016). Similarly, G12V KRAS activation depleted TET1 in a non-malignant lung cell line leading to widespread TSG methylation and cellular transformation (Wu and Brenner, 2014). Assuming that TP53 and KRAS also regulate TET expression in pancreatic cells, each of the three most common PDAC driver mutations would cause a loss in 5'hmC levels and promote TSG methylation. This may help to explain why there are so few PDAC driver mutations, as each mutation would result in the epigenetic inactivation of dozens or hundreds of other tumour suppressors.

SMAD4 may directly activate TET2 transcription, or may influence another factor that promotes TET2. SMAD4 can bind to the distal end of the TET2 gene, close to an enhancer element, however it is not clear whether this enhancer regulates the TET2 gene or not as enhancers can act on promoters hundreds of kb away. This interaction could be elucidated through the use of chromatin conformation capture methods such as 3C, 4C and Hi-C (reviewed by de Wit and de Laat (2012)) which

can determine DNA sequences that are in close physical proximity such as promoters and enhancers. Until such an experiment is done it remains unclear whether TET2 is a direct target of SMAD4 or not.

SMAD4 loss has long been linked to increased tumour aggressiveness, as well as the aggressive basal subtype identified by Moffitt et al. (2015). SMAD4 can also promote the expression of classical-pancreatic associated genes such as E-Cadherin (Muller et al., 2002). In this chapter it was shown that SMAD4 can also promote the expression of GATA6, a key regulator of subtype identity, as well as PDX1. Furthermore, re-expression of SMAD4 in SMAD4-null cell lines decreased the enrichment for squamous associated gene sets, indicating that loss of SMAD4 can promote a more squamous subtype. While TET2 can regulate subtype genes in a similar way, and SMAD4 promotes TET2 expression, it is unclear if the regulation of molecular subtype genes by SMAD4 is dependent on TET2 or not. To help answer this question it would be interesting the measure the effects of SMAD4 inhibition in a TET2 null cell line.

Chapter 7

Discussion

7.1 Overview of Results

7.1.1 5'hmC is a Tumour Suppressor in PDAC

In this study the role of 5'hmC in PDAC was investigated. During pancreatic differentiation 5'hmC is dynamically regulated and accumulates at key differentiation genes (Li et al., 2018). Whole genome IHC staining for 5'hmC in patient tumours showed a dramatic reduction in 5'hmC levels in PDAC. This could be seen at the earliest detectable stages of tumourigenesis, with low grade PanIN lesions displaying a 50% reduction in IHC staining. KRAS G12 mutations are typically considered as the initiating factor in PDAC tumorigenesis due to their presence in 90% of PDAC tumours (Waters and Der, 2018). However, hyperactivation of RAS proteins activates tumour suppressor genes which must be silenced in order to allow cells to continue to proliferate. Therefore, RAS-driven transformation of cells is concomitant with the methylation-associated silencing of these TSGs (Patra, 2008)(Gazin et al., 2007). Interestingly, KRAS inhibits TET1 expression via activation of the ERK pathway, resulting in a drop in 5'hmC and leading to TSG silencing which is critical for the KRAS driven transformation of non-malignant cell lines (Wu and Brenner, 2014).

Therefore, the initial drop in 5'hmC seen in low grade PanIN lesions may be due to KRAS inhibition of TET1. 5'hmC levels continue to drop as PanIN lesions progress and PDAC tumours form, which may be caused by the other TET proteins. Notably, TP53 mutations result in lower TET1 and TET2 expression (Tovy et al., 2017), and it was shown in the previous chapter that loss of SMAD4 correlates with reduced TET2. Therefore, inhibition of TET1 and TET2 may play a role in the decrease in 5'hmC seen during PDAC tumorigenesis.

Next, whole genome 5'hmC distribution was determined by adaption of the oxBS method for FFPE material. This provided site-specific information of 5'hmC in a large cohort of patients for the first time and showed that 5'hmC is being lost at PDAC associated genes. This included TP53 targets, differentiation associated genes and genes that are downregulated in PDAC. Furthermore, patients who maintained the highest level of 5'hmC at these tumour specific differentially hydroxymethylated probes had better disease free survival. Peaks of 5'hmC at the edges of regulatory regions can protect these regions from DNMT activity (Li et al., 2018). In PDAC tumours it was found that loss of 5'hmC at CpG shores inversely correlated with 5'mC at CpG islands, in line with the idea that 5'hmC helps maintain methylation-free regions. Taken together, these results suggest that 5'hmC acts as a tumour suppressor in PDAC by preventing tumour suppressors from 5'mC-mediated silencing.

7.1.2 TET2 Regulates Molecular Subtype Genes

Integrated analysis of 5'mC, 5'hmC and genomic mutations split the PDAC cohort into two clusters with different clinical outcomes. One group was highly aggressive and associated with low 5'hmC, mutations in SMAD4 and TP53 and increased 5'mC at differentiation associated genes. This group showed a high level of overlap with the highly aggressive squamous subtype which is also associated with TP53 mutations, decreased SMAD4 mRNA and silencing of differentiation genes (Bailey et al.,

2016)(Moffitt et al., 2015)(Collisson et al., 2011), suggesting that lower 5'hmC may also be involved in development of a squamous subtype.

The endodermal differentiation marker GATA6 is a marker for classical-pancreatic subtypes, the loss of which allows tumours to become Wnt-autonomous (Seino et al., 2018). GATA6 is silenced by promoter hypermethylation in squamous tumours (Bailey et al., 2016), however the cause of this was previously unknown. In this thesis it was shown that 5'hmC is reduced at the GATA6 gene in squamous tumours, which correlates with increased 5'mC at the GATA6 promoter and decreased mRNA expression. GATA6 mRNA also correlates with TET2 mRNA, but not TET1 or TET3. These results suggest that GATA6 expression is regulated by TET2-mediated 5'hmC. These *in silico* observations were partially validated *in vitro* when TET2 overexpression in two squamous cell lines increased GATA6 expression. However, it was not determined what effect TET2 overexpression had on 5'mC and 5'hmC at the GATA6 gene. Interestingly, TET2 overexpression also increased PDX1 and E-Cadherin, two other classical-pancreatic markers, and decreased N-Cadherin, a squamous marker; consistent with a switch in the molecular subtype of these cells. These results are in line with emerging data that point to the plasticity of molecular subtypes (Candido et al., 2018) and suggest that it is possible to convert squamous tumours to classical-pancreatic.

7.1.3 Pharmacological Targeting of TET2

The differences in survival between tumours with different molecular subtypes suggests that pharmacological switching of squamous to classical-pancreatic could have profound effects on patient outcomes. Given its constant association with classical-pancreatic subtypes (Collisson et al., 2019) and potential role in maintaining these subtypes (Seino et al., 2018), restoration of GATA6 is the most promising candidate for causing a change in subtype. There are no drugs currently available that

directly promote GATA6 expression. However, there are multiple drugs known to target TET2. Two such drugs are metformin and ascorbic acid which promote TET2 stability and activity respectively. In all three squamous cell lines tested combined metformin/AA treatment increased both TET2 and GATA6 levels, indicating that it is possible to restore GATA6 in squamous cell lines pharmacologically. Additionally, administration of metformin/AA to mice with orthotopic squamous PDAC tumours showed that GATA6 can be restored *in vivo*. Critically, *in vitro* experiment were performed using clinically relevant drug concentrations, and *in vivo* tumours recapitulate aspects of human PDAC tumour microenvironment, making it more likely that these results are translatable to clinical settings.

RNA-seq analysis of orthotopic PANC1 tumours showed that GATA6 mRNA was increased in both tumour and stroma cells. Furthermore, Wnt ligands were also inhibited, consistent with the role of GATA6 in maintaining Wnt dependence (Seino et al., 2018). IHC staining showed suppression of nuclear β-catenin following treatment, confirming Wnt pathway inhibition. Thus, GATA6 restoration may prevent squamous tumours from activating their own Wnt signalling, whilst also inhibiting Wnt ligand secretion from the stroma. However, this was only sufficient to cause a partial switch in molecular subtype, as the enrichment for squamous gene sets was reduced but classical-pancreatic gene sets did not increase following treatment. Therefore, additional strategies may be required in order to fully switch molecular subtypes.

Interestingly, Metformin/AA also increased GATA6 levels in a classical-pancreatic cell line. It is still largely unknown how squamous tumours develop but one possibility is that classical-pancreatic tumours become squamous following GATA6 loss. Therefore, by treating tumours with metformin and AA it may be possible to maintain high levels of GATA6 and prevent classical-pancreatic tumours from progressing to squamous. However, further research into the evolution of PDAC molecular subtypes is required to determine if this could indeed be the case.

7.2 Potential Mechanisms of TET2 Loss in PDAC

If TET2 is a major contributor to subtype identity then any factor that causes TET2 loss should also promote a more squamous subtype. Here, some of the possible mechanisms by which TET2 could be lost in PDAC tumours will be explored, as well as the therapeutic interventions that could target them.

7.2.1 Mutations

Mutations in the TET2 gene are rare in PDAC tumours, however other common PDAC mutations may contribute to TET2 loss. In the previous chapter SMAD4 was shown to regulate TET2 expression. Similarly, in ESCs TP53 has been shown to promote TET2, with TP53 mutations leading to loss of 5'hmC and widespread hypermethylation (Tovy et al., 2017). TP53 and SMAD4 are the second and third most common PDAC driver mutations respectively and have long been linked to increased tumour aggressiveness (Grant et al., 2016), as well as more aggressive molecular subtypes (Bailey et al., 2016)(Moffitt et al., 2015). Therefore, TP53 and SMAD4 loss may contribute to progression to a squamous subtype in part through downregulation of TET2 expression. However, assuming that subtypes are plastic and some squamous tumours derive from classical-pancreatic tumours, one clue as to how molecular subtypes arise may come from the mutually exclusive nature of subtype gene sets. This would suggest that at any one time (i.e. when the sample is taken) tumours are either classical-pancreatic or squamous, but not a mix of the two. This would argue against a new subtype arising from a single cell that then outcompetes the rest of the tumour, as would be the case for a genomic mutation, as the tumour would be a mix of both subtypes, at least for a short period of time. Instead, it is more likely that a switch in subtype would be caused by an environmental factor that affects the whole tumour. Notably, the conversion of environmental stimuli to phenotypic

responses is typically mediated through epigenetic changes in gene expression (Tiffon, 2018)(Romani et al., 2015). Therefore, reduced TET2 expression via TP53 and SMAD4 mutations may weaken the epigenetic barrier to subtype switching while the environment pushes these tumours over the edge into a squamous phenotype.

7.2.2 Regulation by other Epigenetic Enzymes

The epigenetic activation or inactivation of a gene is not caused by a single mark, but rather through the concerted regulation of multiple enzymes. For example, sustained gene activation is only achieved when H3K4me3 and H3K79me3 are present at the promoter region in the absence of 5'mC (Cano-Rodriguez et al., 2016). It is therefore unsurprising that epigenetic enzymes would regulate the activity of one another. Histone Acetyltransferase p300 (p300) acetylates multiple histone sites associated with increased transcriptional activation, including histone 3 at lysine 122 and 27 (H3K122ac and H3K27ac) (Ogryzko et al., 1996)(Tropberger et al., 2013). This is antagonized by Histone De-acetylases (HDACs) leading to gene silencing (Chen et al., 2017). These proteins also have numerous other non-histone targets. One examples is P53 which is acetylated by P300 and de-acetylated by HDAC1 leading to increased or decreased activation respectively (Ito et al., 2001)(Luo et al., 2000).Zhang and Shay (2017) found that TET2 is also a target of P300 and HDAC1/2, with acetylation of TET2 causing increased protein stability and activity. HDACs are frequently overexpressed in multiple cancers, including PDAC (Reviewed by Li and Seto (2016)), which could lead to lower TET2 levels. Multiple HDAC inhibitors have now been been developed and are now being used to treat a variety of cancers. However, clinical trials in PDAC patients are still in the early stages (Koutsounas et al., 2013)(Chan et al., 2016). As HDAC1/2 can also inhibit TET2 the use of HDAC inhibitors such as vorinostat may result in TET2 restoration.

7.2.3 Diabetes and Hyperglycemia

If environmental factors are a major driver of subtype identity, then one possibly candidate for development of a squamous subtype that was explored throughout this thesis is diabetes and hyperglycemia. At the time of diagnosis 85% of PDAC patients have hyperglycemia and 50% have diabetes (Permert et al., 1993). Many of these patients do not have a previous history of diabetes (Mekan et al., 2006) and in many cases hyperglycemia is the only indication that there is a tumour present prior to diagnosis (Sah et al., 2019). Despite this, the role of hyperglycemia in PDAC remains unclear. Under hyperglycemic conditions TET2 becomes destabilized due to inhibition of AMPK (Wu et al., 2018). The nutrient sensor mTOR on the other hand becomes hyperactivated during nutrient overload to promote cell growth (Yoon, 2017) and suppress TET2 expression (Liu et al., 2013). Therefore, hyperglycemia may inhibit TET2 levels through multiple mechanisms leading to a more aggressive tumour.

Glucose levels can be effectively modulated using the anti-diabetic drug metformin. The results in this thesis suggest that metformin may protect against squamous subtypes by increasing TET2 and GATA6 levels. A meta-analysis on the effects of metformin on pancreatic cancer patient survival, including 9 retrospective cohorts and 2 clinical trials, found that metformin was significantly associated with better survival in resected or locally advanced tumours (HR = 0.79 and 0.68 respectively), but not in metastatic tumours (Li et al., 2017). As most PDAC tumours are metastatic at diagnosis, metformin treatment alone in combination with conventional treatment is not sufficient. One method to increase the effectiveness of metformin may be the addition of ascorbic acid which was shown to have a combinatorial effect on TET2 and GATA6. In three squamous cell lines the combination of metformin and AA was able to increase GATA6 and TET2 levels at physiological levels. A similar result was seen *in vivo*, resulting in Wnt pathway inhibition and decreased tumour prolifera-

tion. Thus, this combination may increase the effectiveness of metformin in PDAC patients. If given to a patient with a squamous tumour it may make the tumour more classical-pancreatic and thereby less aggressive. Alternatively, as metformin/AA can also increase GATA6 in classical-pancreatic cells it may prevent classical-pancreatic tumours from losing GATA6 and becoming squamous.

Various mTOR inhibits have also been developed that may give similar effects to metformin on TET2 and GATA6 levels. Most notable of these is rapamycin, for which mTOR was named, as well as its analogs such as everolimus. These drugs are well tolerated by patients but have failed to be effective in three phase 2 metastatic PDAC trials (Reviewed by Iriana et al. (2016). Rapamycin can increase TET2 in cultured smooth muscle cells (Liu et al., 2013). Therefore, the effects of rapamycin may also increase when used in combination with metformin or AA. However, a common side effect of rapamycin and everolimus in PDAC patients is hyperglycemia. As hyperglycemia inhibits TET2 it is unclear if the results seen *in vitro* will be translatable to in vivo tumours, which could partially explain why these drugs failed in PDAC patients.

7.2.4 Hypoxia

Another environmental factor that may contribute to TET2 loss is oxygen levels. The dense tumour stroma characteristic of PDAC exerts significant interstitial pressure of the tumour vasculature, causing blood vessels to compress and leading to an oxygen poor environment referred to as hypoxia (Provenzano et al., 2012). This activates the Hypoxia-inducible Factors (HIFs) which drive pro-survival signalling (Laitala and Erler, 2018), as well as increased stemness and chemoresistance (McGinn et al., 2017). Under normal oxygen levels the alpha-subunits of HIFs (HIF-1/2α) are hydroxylated by Prolyl Hydroxylase Domain-Containing proteins (PHDs) (Jaakkola et al., 2001) and Factor Inhibiting HIF-1 (FIH-1) (Mahon et al., 2001), leading to HIF-1/2α degra-

dation. PHDs and FIH-1 belong to the Fe^{2+} and α-KG dependent dioxygenase family and require oxygen for their activity. Thus, in the absence of oxygen these enzymes become inactive and HIF1/2α accumulate.

PHDs and FIH-1 belong to the same family as the TET enzymes, all of which require Fe^{2+}, α-KG and oxygen. Therefore, under low oxygen conditions TET activity is also lost (Thienpont et al., 2016)(Burr et al., 2018). Additionally, TET2 is suppressed by HIF-1α independently of this as knockdown of HIF-1α increases TET2 mRNA and protein expression (Fischer and Miles, 2017). Thus, hypoxia promotes loss of TETs through multiple mechanisms. Interestingly, as ascorbic acid is a co-factor for the Fe^{2+} and α-KG dependent dioxygenase family AA treatment can suppress HIF-1α by increasing PHD and FIH-1 activity (Kuiper and Vissers, 2014)(Miles et al., 2015) which could also cause TET2 levels to increase. HIF-1α levels were examined following AA treatment in PSN1 cells (n=1) but no changes could be seen (see appendix), possibly because cells are cultured in atmospheric (20%) oxygen concentrations. However, it is unclear if this is also the case in orthotopic tumours which could have developed a hypoxic core. To determine if this is the case a marker of hypoxia such as Carbonic Anhydrase 9 (CA9) could be examined.

Metformin is another drug that can influence the hypoxic response by suppressing mitochondrial oxygen consumption in an AMPK-independent manner. This decreases the oxygen consumption of cells leading to increased intracellular oxygen levels and HIF-1α degradation (Zhou et al., 2016). Metformin has been used to suppress hypoxia-induced migration and drug resistance in multiple tumour types, including PDAC (Fujita et al., 2019)(Ye et al., 2018)(Suzuki et al., 2019). Therefore, metformin may also increase TET2 levels by suppressing hypoxia and HIF-1α. In this thesis it was shown that metformin-induced increases in TET2 and GATA6 are AMPK-dependent in PSN1 in culture so this mechanism also does not appear to be active in the cultured cells used.

Finally, various HDAC inhibitors has been shown to suppress HIF-1α activity, possibly through transcriptional inhibition (Kim et al., 2007)(Hutt et al., 2014), providing another potential mechanism for drugs such as vorinostat to increase TET2 levels. An overview of the potential contributing factors to TET2 loss in PDAC is shown in Figure 7.1.

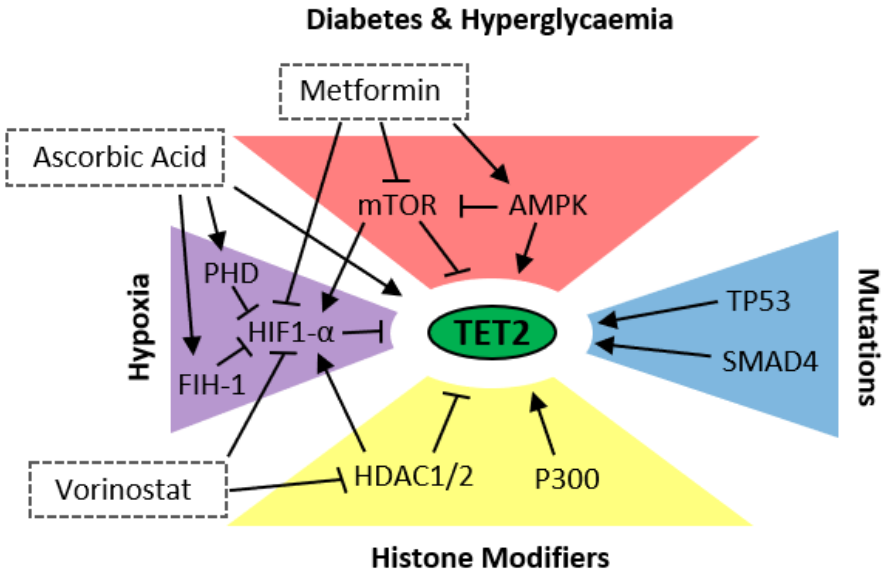


Figure 7.1: **Potential Mechanisms of TET2 Loss in PDAC**. Schematic showing factors contributing to TET2 loss in PDAC. Candidate drugs to target these interactions are shown in grey boxes. Only metformin and ascorbic acid were investigated in this thesis these but multiple other drugs may have similar effects on TET2 and may have synergistic effects when used in combination.

It should be noted that all of the factors described above occur in a wide range of PDAC tumours of both subtypes. For example, while TP53 mutations are enriched in squamous tumours, TP53 is the second most common PDAC mutation and also occurs in many classical-pancreatic tumours (Bailey et al., 2016). Similarly, hyperglycemia is present in 85% of PDAC tumours at diagnosis (Permert et al., 1993). This indicates that any one factor is not an absolute determinant of subtype identity. However, TP53/SMAD4 mutations, HDAC overexpression, hyperglycemia and hypoxia are each known to increase tumour aggressiveness. Thus, it may be a com-

bination of these factors, each contributing to TET2 loss, that ultimately determines whether a tumour becomes squamous or not.

7.3 Remaining Aspects and Future Work

7.3.1 Further Validation of the oxBS Method in FFPE

One of the outcomes of this thesis is an FFPE optimized protocol for the oxBS method. To the best of my knowledge this is the first time a method for investigating 5'hmC distribution genome-wide has been utilized in FFPE material. There are numerous pieces of evidence that suggest that this optimization was successful and the results obtained are due to 5'hmC and not background noise. However, to fully validate this method the results obtained in FFPE tissue should be compared to those obtained from fresh frozen tissue.

7.3.2 Genome-wide 5'hmC Changes in Treated Cells

In PDAC tumours there are differences in the distribution of 5'hmC and 5'mC at the GATA6 gene between squamous and classical-pancreatic samples. Overexpression of TET2 increased GATA6 expression showing that TET2 regulates GATA6, however it was not shown that this was due to TET2 de-methylating the GATA6 gene. To test if this is indeed the case 5'hmC and 5'mC distribution changes should be examined following TET2 overexpression. Similar experiments should also be performed in cells treated with metformin and ascorbic acid.

7.3.3 Determining the Extent of TET2 Involvement in Subtype Changes

In this thesis it was shown that TET2 regulates subtype specific genes. Furthermore, treatment with metformin and AA, which both promote TET2, can increase GATA6. However, both metformin and AA regulate multiple signalling pathways so it remains unclear to what extent TET2 is responsible for restoring GATA6 expression. An answer to this question was attempted multiple times by knocking down TET2 using siRNAs and then testing the effects of drugs. However, in all cases the TET2 knockdown did not work. This was not due to transfection efficiency as SMAD4 was successfully knocked down in some of the cell lines used. It may be that TET2 knockdown cells cannot compete with the surrounding untransfected cells and their signal becomes lost. Therefore, a TET2 knockout cell line may need to be generated instead in order to isolate a population of cells with complete TET2 loss.

7.4 Conclusion

In conclusion, the results in this thesis identify TET2 as a regulator of PDAC molecular subtypes that acts upstream of GATA6. From these results the following model is proposed: In classical-pancreatic cells TET2 maintains the active state of the GATA6 promoter by preventing DNMT-mediated silencing. Because of the high levels of GATA6 these cells are dependent on the surrounding stroma to supply them with Wnt ligands in order to promote Wnt pathway activation. In squamous cells TET2 becomes lost and DNMT is able to methylate the GATA6 promoter. In the absence of GATA6 squamous cells are able to activate their own Wnt signalling and become independent of the stroma, allowing the tumour cells to spread throughout the body (Fig. 7.2). Restoring TET2 using metformin and ascorbic acid can partially reverse this process by increasing GATA6 expression, thereby causing squamous tumours to

become more classical-pancreatic in nature.

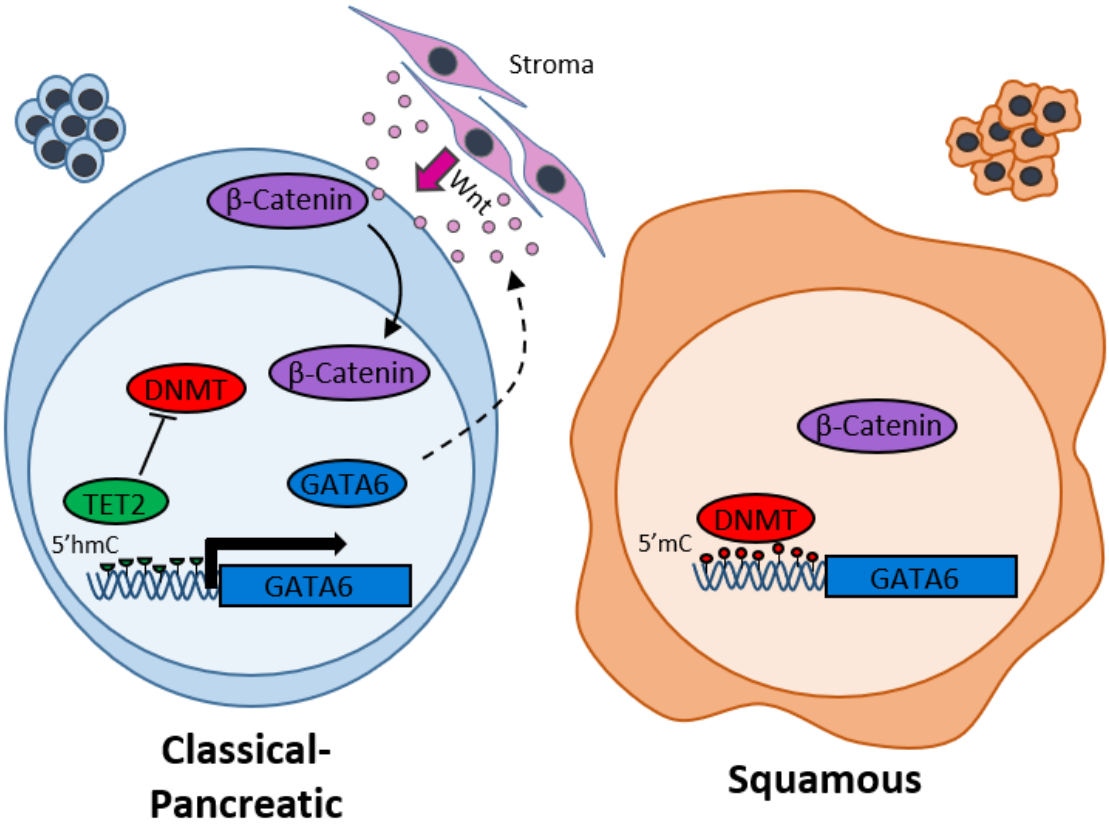


Figure 7.2: **Model for GATA6 Loss in Squamous Tumours**. Schematic showing proposed model for GATA6 loss. In classical-pancreatic tumours the GATA6 gene is hydroxymethylated by TET2, preventing DNMT-mediated silencing. GATA6 causes tumour cells to be dependent on secreted Wnt ligands from stroma cells through an unknown mechanism, which promotes nuclear β-catenin. In squamous cells TET2 is lost, resulting in loss of GATA6 expression. GATA6-null cells can activate their own Wnt signalling in the absence of secreted stromal ligands.

Appendix A

Appendix

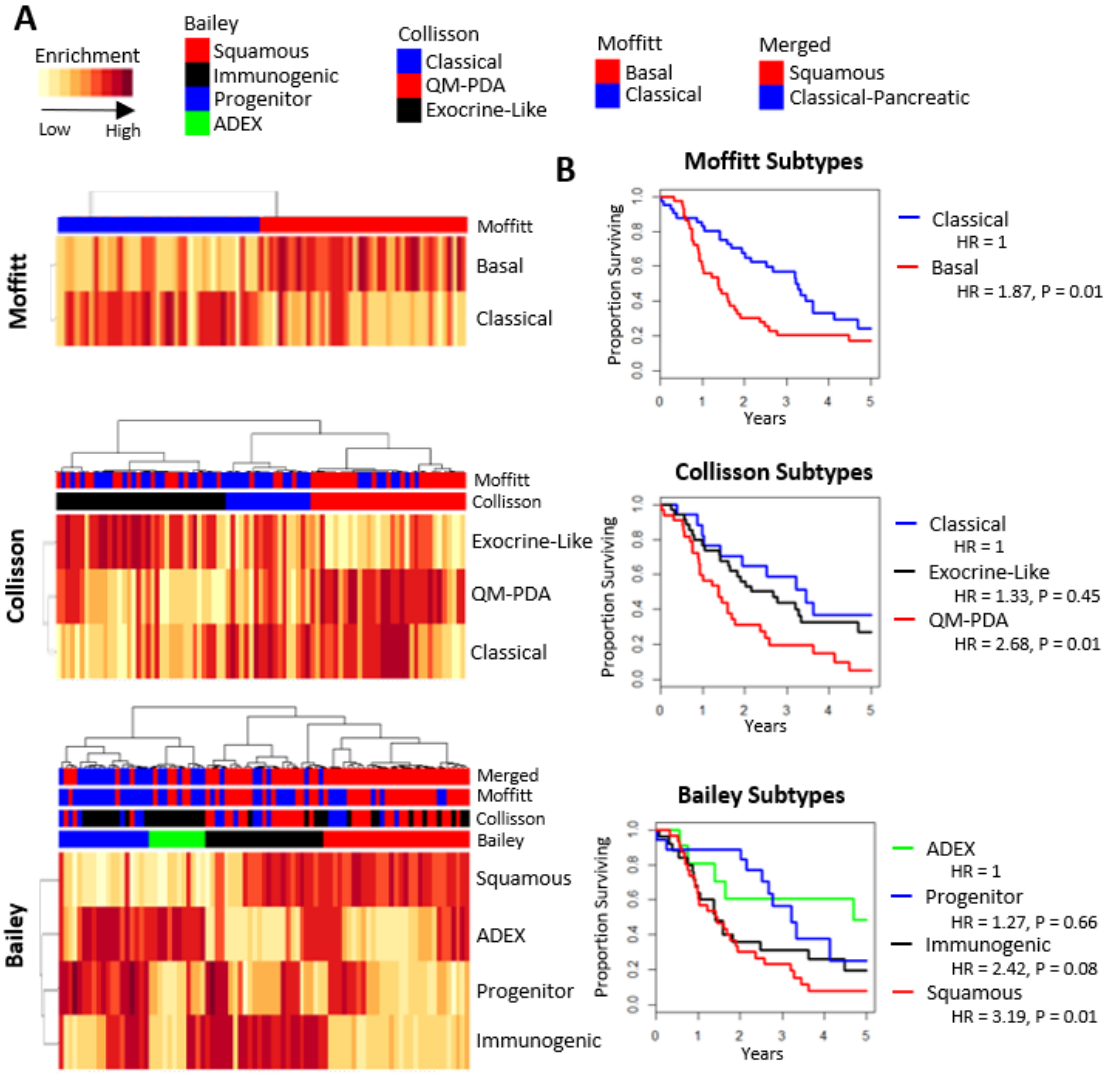


Figure A.1: **Molecular Subtyping of the Oxford Cohort**. **A**) Molecular subtyping of tumours using gene sets defined by Moffitt, Collisson and Bailey. Relative enrichment of each gene set was determined using GSVA. Subtype for each patient is shown in the top panel of each heatmap. B) Kaplan-Meier curves for patients of each subtype. In all cases patients with the more aggressive subtype (Basal, QM-PDA and Squamous) had a worse prognosis

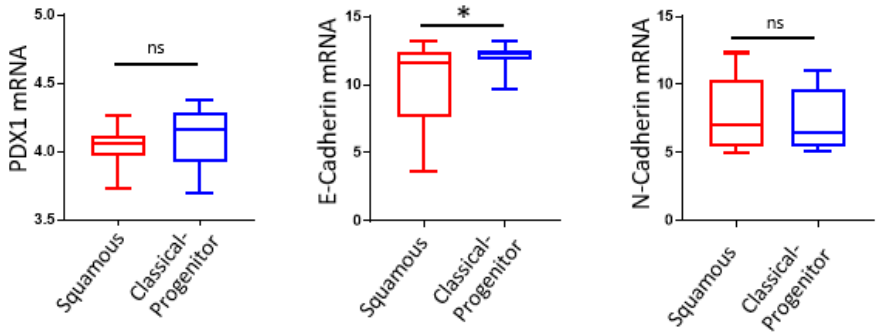


Figure A.2: **mRNA Expression of Subtype Markers in PDAC Cell Lines**. Expression of different subtype markers in squamous and classical-progenitor cell lines. PDX1 and E-Cadherin have been reported as having low expression in squamous tumours, while N-Cadherin is higher. In CCLE Cell lines only E-Cadherin was significantly differentially expressed between subtypes.

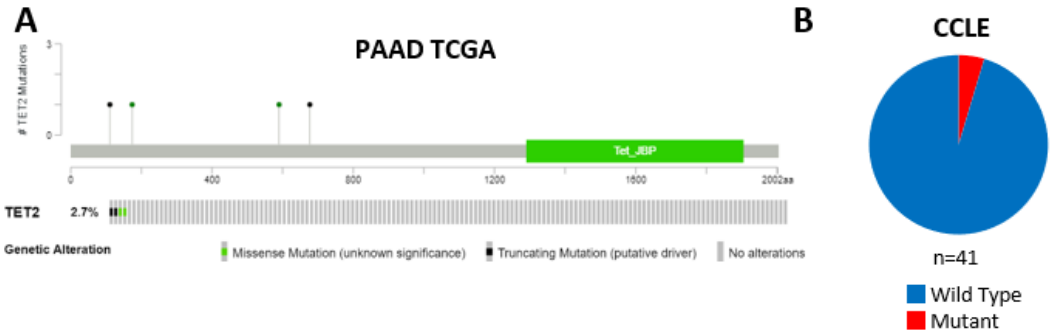


Figure A.3: **TET2 Mutations in PDAC**. **A**) Oncoprint showing mutations in the TET2 gene within the PAAD TCGA cohort. Downloaded from cBioPortal. **B**) Proportion of CCLE PDAC cell lines with TET2 mutations.

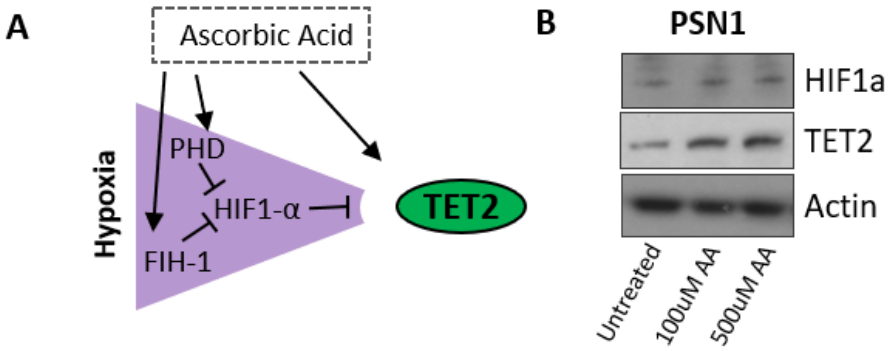


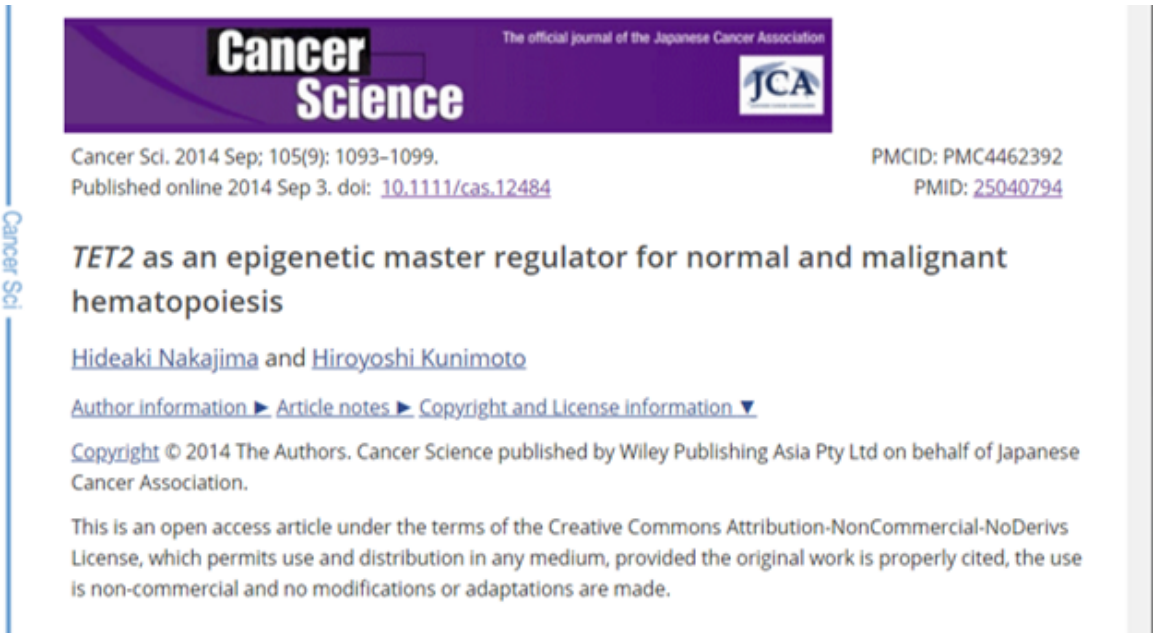
Figure A.4: **Ascorbic Acid Treatment Does not Alter HIF-1α *in vitro*.** **A**) Schematic showing action of ascorbic acid on TET2 and HIF-1α regulators. **B**) Western blot of PSN1 cells treated with 100uM and 500uM AA. No change in HIF-1α levels are seen (n=1).

Appendix B

Copyright Licences



Figure B.1: Copyright notice: Zaidi et al, 2010. Architectural Epigenetics: Mitotic Retention of Mammalian Transcriptional Regulatory Information.



Cancer Science
The official journal of the Japanese Cancer Association
JCA

Cancer Sci. 2014 Sep; 105(9): 1093–1099.
Published online 2014 Sep 3. doi: 10.1111/cas.12484

PMCID: PMC4462392
PMID: 25040794

Cancer Sci

TET2 as an epigenetic master regulator for normal and malignant hematopoiesis

Hideaki Nakajima and Hiroyoshi Kunimoto

Author information ▶ Article notes ▶ Copyright and License information ▼

Copyright © 2014 The Authors. Cancer Science published by Wiley Publishing Asia Pty Ltd on behalf of Japanese Cancer Association.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Figure B.2: Copyright notice: Nakajima and Kunimoto, 2014. TET2 as an epigenetic mastor regulator for normal and malignant hematopoiesis.

Bibliography

D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–74, 2011. ISSN 0092-8674. doi: 10.1016/j.cell.2011.02.013. 1, 2

E. Clayton, D. P. Doupe, A. M. Klein, D. J. Winton, B. D. Simons, and P. H. Jones. A single type of progenitor cell maintains normal epidermis. *Nature*, 446(7132): 185–9, 2007. ISSN 0028-0836. doi: 10.1038/nature05574. 3

N. Barker, J. H. van Es, J. Kuipers, P. Kujala, M. van den Born, M. Cozijnsen, A. Haegebarth, J. Korving, H. Begthel, P. J. Peters, and H. Clevers. Identification of stem cells in small intestine and colon by marker gene lgr5. *Nature*, 449(7165): 1003–7, 2007. ISSN 0028-0836. doi: 10.1038/nature06196. 3

H. Tian, B. Biehs, S. Warming, K. G. Leong, L. Rangell, O. D. Klein, and F. J. de Sauvage. A reserve stem cell population in small intestine renders lgr5-positive cells dispensable. *Nature*, 478(7368):255–9, 2011. ISSN 0028-0836. doi: 10.1038/nature10408. 3

J. H. van Es, T. Sato, M. van de Wetering, A. Lyubimova, A. N. Yee Nee, A. Gregorieff, N. Sasaki, L. Zeinstra, M. van den Born, J. Korving, A. C. M. Martens, N. Barker, A. van Oudenaarden, and H. Clevers. Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nat Cell Biol*, 14(10):1099–1104, 2012. ISSN 1465-7392. doi: 10.1038/ncb2581. 3

E. Batlle and H. Clevers. Cancer stem cells revisited. *Nat Med*, 23(10):1124–1134, 2017. ISSN 1078-8956. doi: 10.1038/nm.4409. 3, 4

A. Fumagalli, J. Drost, S. J. Suijkerbuijk, R. van Boxtel, J. de Ligt, G. J. Offerhaus, H. Begthel, E. Beerling, E. H. Tan, O. J. Sansom, E. Cuppen, H. Clevers, and J. van Rheenen. Genetic dissection of colorectal cancer progression by orthotopic transplantation of engineered cancer organoids. *Proc Natl Acad Sci U S A*, 114 (12):E2357–e2364, 2017. ISSN 0027-8424. doi: 10.1073/pnas.1701219114. 4

M. Fujii, M. Shimokawa, S. Date, A. Takano, M. Matano, K. Nanki, Y. Ohta, K. Toshimitsu, Y. Nakazato, K. Kawasaki, T. Uraoka, T. Watanabe, T. Kanai, and T. Sato. A colorectal tumor organoid library demonstrates progressive loss of niche factor requirements during tumorigenesis. *Cell Stem Cell*, 18(6):827–38, 2016. ISSN 1875-9777. doi: 10.1016/j.stem.2016.04.003. 4

G. Driessens, B. Beck, A. Caauwe, B. D. Simons, and C. Blanpain. Defining the mode of tumour growth by clonal analysis. *Nature*, 488(7412):527–30, 2012. ISSN 0028-0836. doi: 10.1038/nature11344. 4

R. L. Pictet, W. R. Clark, R. H. Williams, and W. J. Rutter. An ultrastructural analysis of the developing embryonic pancreas. *Dev Biol*, 29(4):436–67, 1972. ISSN 0012-1606 (Print) 0012-1606. doi: 10.1016/0012-1606(72)90083-8. 4

J. Jonsson, L. Carlsson, T. Edlund, and H. Edlund. Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature*, 371(6498):606–9, 1994. ISSN 0028-0836 (Print) 0028-0836. doi: 10.1038/371606a0. 4

G. Kloppel and J. Luttges. Who-classification 2000: exocrine pancreatic tumors. *Verh Dtsch Ges Pathol*, 85:219–28, 2001. ISSN 0070-4113 (Print) 0070-4113. 5

R. M. M. Ferreira, R. Sancho, H. A. Messal, E. Nye, B. Spencer-Dene, R. K. Stone, G. Stamp, I. Rosewell, A. Quaglia, and A. Behrens. Duct- and acinar-derived pancreatic ductal adenocarcinomas show distinct tumor progression and marker expression. *Cell Rep*, 21(4):966–978, 2017. doi: 10.1016/j.celrep.2017.09.093. 5, 12

JP. Morris, SC. Wang, and M. Hebrok. Kras, hedgehog, wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. *Nat Rev Cancer*, 10(10), 2010. ISSN 683-95. doi: 10.1038/nrc2899. 6

J. L. Humphris, A. M. Patch, K. Nones, P. J. Bailey, A. L. Johns, S. McKay, D. K. Chang, D. K. Miller, M. Pajic, K. S. Kassahn, M. C. Quinn, T. J. Bruxner, A. N. Christ, I. Harliwong, S. Idrisoglu, S. Manning, C. Nourse, E. Nourbakhsh, A. Stone, P. J. Wilson, M. Anderson, J. L. Fink, O. Holmes, S. Kazakoff, C. Leonard, F. Newell, N. Waddell, S. Wood, R. S. Mead, Q. Xu, J. Wu, M. Pinese, M. J. Cowley, M. D. Jones, A. M. Nagrial, V. T. Chin, L. A. Chantrill, A. Mawson, A. Chou, C. J. Scarlett, A. V. Pinho, I. Rooman, M. Giry-Laterriere, J. S. Samra, J. G. Kench, N. D. Merrett, C. W. Toon, K. Epari, N. Q. Nguyen, A. Barbour, N. Zeps, N. B. Jamieson, C. J. McKay, C. R. Carter, E. J. Dickson, J. S. Graham, F. Duthie, K. Oien, J. Hair, J. P. Morton, O. J. Sansom, R. Grutzmann, R. H. Hruban, A. Maitra, C. A. Iacobuzio-Donahue, R. D. Schulick, C. L. Wolfgang, R. A. Morgan, R. T. Lawlor, B. Rusev, V. Corbo, R. Salvia, I. Cataldo, G. Tortora, M. A. Tempero, O. Hofmann, J. R. Eshleman, C. Pilarsky, A. Scarpa, E. A. Musgrove, A. J. Gill, J. V. Pearson, S. M. Grimmond, N. Waddell, and A. V. Biankin. Hypermutation in pancreatic cancer. *Gastroenterology*, 152(1):68–74.e2, 2017. ISSN 0016-5085. doi: 10.1053/j.gastro.2016.09.060. 5

V. T. Smit, A. J. Boot, A. M. Smits, G. J. Fleuren, C. J. Cornelisse, and J. L. Bos. Kras codon 12 mutations occur very frequently in pancreatic adenocarcinomas. *Nucleic Acids Res*, 16(16):7773–82, 1988. ISSN 0305-1048 (Print) 0305-1048. doi: 10.1093/nar/16.16.7773. 5

A. J. Aguirre, J. A. Nowak, N. D. Camarda, R. A. Moffitt, A. A. Ghazani, M. Hazar-Rethinam, S. Raghavan, J. Kim, L. K. Brais, D. Ragon, M. W. Welch, E. Reilly,

D. McCabe, L. Marini, K. Anderka, K. Helvie, N. Oliver, A. Babic, A. Da Silva, B. Nadres, E. E. Van Seventer, H. A. Shahzade, J. P. St Pierre, K. P. Burke, T. Clancy, J. M. Cleary, L. A. Doyle, K. Jajoo, N. J. McCleary, J. A. Meyerhardt, J. E. Murphy, K. Ng, A. K. Patel, K. Perez, M. H. Rosenthal, D. A. Rubinson, M. Ryou, G. I. Shapiro, E. Sicinska, S. G. Silverman, R. J. Nagy, R. B. Lanman, D. Knoerzer, D. J. Welsch, M. B. Yurgelun, C. S. Fuchs, L. A. Garraway, G. Getz, J. L. Hornick, B. E. Johnson, M. H. Kulke, R. J. Mayer, J. W. Miller, P. B. Shyn, D. A. Tuveson, N. Wagle, J. J. Yeh, W. C. Hahn, R. B. Corcoran, S. L. Carter, and B. M. Wolpin. Real-time genomic characterization of advanced pancreatic cancer to enable precision medicine. *Cancer Discov*, 8(9):1096–1111, 2018. ISSN 2159-8274. doi: 10.1158/2159-8290.Cd-18-0275. 5, 116

J. G. Reiter, A. P. Makohon-Moore, J. M. Gerold, A. Heyde, M. A. Attiyeh, Z. A. Kohutek, C. J. Tokheim, A. Brown, R. M. DeBlasio, J. Niyazov, A. Zucker, R. Karchin, K. W. Kinzler, C. A. Iacobuzio-Donahue, B. Vogelstein, and M. A. Nowak. Minimal functional driver gene heterogeneity among untreated metastases. *Science*, 361 (6406):1033–1037, 2018. ISSN 0036-8075. doi: 10.1126/science.aat7171. 5, 18, 50

Z. A. Rasheed, W. Matsui, and A. Maitra. *Pathology of pancreatic stroma in PDAC*. Transworld Research Network Transworld Research Network., Trivandrum (India), 2012. 7, 99

A. Maitra and R. H. Hruban. Pancreatic cancer. *Annu Rev Pathol*, 3:157–88, 2008. ISSN 1553-4006 (Print) 1553-4006. doi: 10.1146/annurev.pathmechdis.3.121806.154305. 7, 99

D. Thomas and P. Radhakrishnan. Tumor-stromal crosstalk in pancreatic cancer and tissue fibrosis. *Mol Cancer*, 18(1):14, 2019. ISSN 1476-4598. doi: 10.1186/s12943-018-0927-5. 7, 99

E. Lonardo, J. Frias-Aldeguer, P. C. Hermann, and C. Heeschen. Pancreatic stellate cells form a niche for cancer stem cells and promote their self-renewal and invasiveness. *Cell Cycle*, 11(7):1282–90, 2012. ISSN 1551-4005. doi: 10.4161/cc.19679. 7

T. Seino, S. Kawasaki, M. Shimokawa, H. Tamagawa, K. Toshimitsu, M. Fujii, Y. Ohta, M. Matano, K. Nanki, K. Kawasaki, S. Takahashi, S. Sugimoto, E. Iwasaki, J. Takagi, T. Itoi, M. Kitago, Y. Kitagawa, T. Kanai, and T. Sato. Human pancreatic tumor organoids reveal loss of stem cell niche factor dependence during disease progression. *Cell Stem Cell*, 22(3):454–467.e6, 2018. ISSN 1875-9777. doi: 10.1016/j.stem.2017.12.009. 7, 16, 67, 99, 107, 111, 114, 131, 132

C. Liang, S. Shi, Q. Meng, D. Liang, S. Ji, B. Zhang, Y. Qin, J. Xu, Q. Ni, and X. Yu. Complex roles of the stroma in the intrinsic resistance to gemcitabine in pancreatic cancer: where we are and where we are going. *Exp Mol Med*, 49(12): e406, 2017. ISSN 1226-3613. doi: 10.1038/emm.2017.255. 7

P. P. Provenzano, C. Cuevas, A. E. Chang, V. K. Goel, D. D. Von Hoff, and S. R. Hingorani. Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer Cell*, 21(3):418–29, 2012. ISSN 1535-6108. doi: 10.1016/j.ccr.2012.01.007. 7, 136

F. Sanchez-Vega, M. Mina, J. Armenia, W. K. Chatila, A. Luna, K. C. La, S. Dimitriadoy, D. L. Liu, H. S. Kantheti, S. Saghafinia, D. Chakravarty, F. Daian, Q. Gao, M. H. Bailey, W. W. Liang, S. M. Foltz, I. Shmulevich, L. Ding, Z. Heins, A. Ochoa, B. Gross, J. Gao, H. Zhang, R. Kundra, C. Kandoth, I. Bahceci, L. Dervishi, U. Dogrusoz, W. Zhou, H. Shen, P. W. Laird, G. P. Way, C. S. Greene, H. Liang, Y. Xiao, C. Wang, A. Iavarone, A. H. Berger, T. G. Bivona, A. J. Lazar, G. D. Hammer, T. Giordano, L. N. Kwong, G. McArthur, C. Huang, A. D. Tward, M. J. Frederick, F. McCormick, M. Meyerson, E. M. Van Allen, A. D. Cherniack, G. Ciriello, C. Sander, and N. Schultz. Oncogenic signaling pathways in the cancer genome atlas. *Cell*, 173(2):321–337.e10, 2018. ISSN 0092-8674. doi: 10.1016/j.cell.2018.03.035. 7

L. Ding, M. H. Bailey, E. Porta-Pardo, V. Thorsson, A. Colaprico, D. Bertrand, D. L. Gibbs, A. Weerasinghe, K. L. Huang, C. Tokheim, I. Cortes-Ciriano, R. Jayasinghe, F. Chen, L. Yu, S. Sun, C. Olsen, J. Kim, A. M. Taylor, A. D. Cherniack, R. Akbani, C. Suphavilai, N. Nagarajan, J. M. Stuart, G. B. Mills, M. A. Wyczalkowski, B. G. Vincent, C. M. Hutter, J. C. Zenklusen, K. A. Hoadley, M. C. Wendl, L. Shmulevich, A. J. Lazar, D. A. Wheeler, and G. Getz. Perspective on oncogenic processes at the end of the beginning of cancer genomics. *Cell*, 173(2):305–320.e10, 2018. ISSN 0092-8674. doi: 10.1016/j.cell.2018.03.033. 8

K. A. Hoadley, C. Yau, T. Hinoue, D. M. Wolf, A. J. Lazar, E. Drill, R. Shen, A. M. Taylor, A. D. Cherniack, V. Thorsson, R. Akbani, R. Bowlby, C. K. Wong, M. Wiznerowicz, F. Sanchez-Vega, A. G. Robertson, B. G. Schneider, M. S. Lawrence, H. Noushmehr, T. M. Malta, J. M. Stuart, C. C. Benz, and P. W. Laird. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell*, 173(2):291–304.e6, 2018. ISSN 0092-8674. doi: 10.1016/j.cell.2018.03.022. 8

E. A. Collisson, A. Sadanandam, P. Olson, W. J. Gibb, M. Truitt, S. Gu, J. Cooc, J. Weinkle, G. E. Kim, L. Jakkula, H. S. Feiler, A. H. Ko, A. B. Olshen, K. L. Danenberg, M. A. Tempero, P. T. Spellman, D. Hanahan, and J. W. Gray. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med*, 17(4):500–3, 2011. ISSN 1078-8956. doi: 10.1038/nm.2344. 8, 57, 67, 79, 80, 131

P. Bailey, D. K. Chang, K. Nones, A. L. Johns, A. M. Patch, M. C. Gingras, D. K. Miller, A. N. Christ, T. J. Bruxner, M. C. Quinn, C. Nourse, L. C. Murtaugh, I. Harliwong, S. Idrisoglu, S. Manning, E. Nourbakhsh, S. Wani, L. Fink, O. Holmes, V. Chin, M. J. Anderson, S. Kazakoff, C. Leonard, F. Newell, N. Waddell, S. Wood, Q. Xu, P. J. Wilson, N. Cloonan, K. S. Kassahn, D. Taylor, K. Quek,

A. Robertson, L. Pantano, L. Mincarelli, L. N. Sanchez, L. Evers, J. Wu, M. Pinese, M. J. Cowley, M. D. Jones, E. K. Colvin, A. M. Nagrial, E. S. Humphrey, L. A. Chantrill, A. Mawson, J. Humphris, A. Chou, M. Pajic, C. J. Scarlett, A. V. Pinho, M. Giry-Laterriere, I. Rooman, J. S. Samra, J. G. Kench, J. A. Lovell, N. D. Merrett, C. W. Toon, K. Epari, N. Q. Nguyen, A. Barbour, N. Zeps, K. Moran-Jones, N. B. Jamieson, J. S. Graham, F. Duthie, K. Oien, J. Hair, R. Grutzmann, A. Maitra, C. A. Iacobuzio-Donahue, C. L. Wolfgang, R. A. Morgan, R. T. Lawlor, V. Corbo, C. Bassi, B. Rusev, P. Capelli, R. Salvia, G. Tortora, D. Mukhopadhyay, G. M. Petersen, D. M. Munzy, W. E. Fisher, S. A. Karim, J. R. Eshleman, R. H. Hruban, C. Pilarsky, J. P. Morton, O. J. Sansom, A. Scarpa, E. A. Musgrove, U. M. Bailey, O. Hofmann, R. L. Sutherland, D. A. Wheeler, A. J. Gill, R. A. Gibbs, J. V. Pearson, N. Waddell, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*, 531(7592):47–52, 2016. ISSN 0028-0836. doi: 10.1038/nature16965. 9, 10, 18, 39, 57, 67, 81, 127, 130, 131, 133, 138

E. A. Collisson, P. Bailey, D. K. Chang, and A. V. Biankin. Molecular subtypes of pancreatic cancer. *Nat Rev Gastroenterol Hepatol*, 16(4):207–220, 2019. ISSN 1759-5045. doi: 10.1038/s41575-019-0109-y. 9, 10, 11, 50, 68, 131

R. A. Moffitt, R. Marayati, E. L. Flate, K. E. Volmar, S. G. Loeza, K. A. Hoadley, N. U. Rashid, L. A. Williams, S. C. Eaton, A. H. Chung, J. K. Smyla, J. M. Anderson, H. J. Kim, D. J. Bentrem, M. S. Talamonti, C. A. Iacobuzio-Donahue, M. A. Hollingsworth, and J. J. Yeh. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet*, 47(10):1168–78, 2015. ISSN 1061-4036. doi: 10.1038/ng.3398. 9, 39, 57, 67, 122, 128, 131, 133

K. A. Hoadley, C. Yau, D. M. Wolf, A. D. Cherniack, D. Tamborero, S. Ng, M. D. M. Leiserson, B. Niu, M. D. McLellan, V. Uzunangelov, J. Zhang, C. Kandoth, R. Akbani, H. Shen, L. Omberg, A. Chu, A. A. Margolin, L. J. Van't Veer, N. Lopez-Bigas, P. W. Laird, B. J. Raphael, L. Ding, A. G. Robertson, L. A. Byers, G. B. Mills, J. N. Weinstein, C. Van Waes, Z. Chen, E. A. Collisson, C. C. Benz, C. M. Perou, and J. M. Stuart. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell*, 158(4):929–944, 2014. ISSN 0092-8674. doi: 10.1016/j.cell.2014.06.049. 10

J. B. Candido, J. P. Morton, P. Bailey, A. D. Campbell, S. A. Karim, T. Jamieson, L. Lapienyte, A. Gopinathan, W. Clark, E. J. McGhee, J. Wang, M. Escorcio-Correia, R. Zollinger, R. Roshani, L. Drew, L. Rishi, R. Arkell, T. R. J. Evans, C. Nixon, D. I. Jodrell, R. W. Wilkinson, A. V. Biankin, S. T. Barry, F. R. Balkwill, and O. J. Sansom. Csf1r(+) macrophages sustain pancreatic tumor growth through t cell suppression and maintenance of key gene programs that define the squamous subtype. *Cell Rep*, 23(5):1448–1460, 2018. doi: 10.1016/j.celrep.2018.03.131. 12, 131

K. L. Aung, S. E. Fischer, R. E. Denroche, G. H. Jang, A. Dodd, S. Creighton, B. Southwood, S. B. Liang, D. Chadwick, A. Zhang, G. M. O'Kane, H. Albaba,

S. Moura, R. C. Grant, J. K. Miller, F. Mbabaali, D. Pasternack, I. M. Lungu, J. M. S. Bartlett, S. Ghai, M. Lemire, S. Holter, A. A. Connor, R. A. Moffitt, J. J. Yeh, L. Timms, P. M. Krzyzanowski, N. Dhani, D. Hedley, F. Notta, J. M. Wilson, M. J. Moore, S. Gallinger, and J. J. Knox. Genomics-driven precision medicine for advanced pancreatic cancer: Early results from the compass trial. *Clin Cancer Res*, 24(6):1344–1354, 2018. ISSN 1078-0432 (Print) 1078-0432. doi: 10.1158/1078-0432.Ccr-17-2994. 13, 75, 81

R. S. Viger, S. M. Guittot, M. Anttonen, D. B. Wilson, and M. Heikinheimo. Role of the gata family of transcription factors in endocrine development, function, and disease. *Mol Endocrinol*, 22(4):781–98, 2008. ISSN 0888-8809 (Print) 0888-8809. doi: 10.1210/me.2007-0513. 13

C. T. Kuo, E. E. Morrisey, R. Anandappa, K. Sigrist, M. M. Lu, M. S. Parmacek, C. Soudais, and J. M. Leiden. Gata4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes Dev*, 11(8):1048–60, 1997. ISSN 0890-9369 (Print) 0890-9369. doi: 10.1101/gad.11.8.1048. 13

R. K. Patient and J. D. McGhee. The gata family (vertebrates and invertebrates). *Curr Opin Genet Dev*, 12(4):416–22, 2002. ISSN 0959-437X (Print) 0959-437x. 13

M. J. Weiss and S. H. Orkin. Gata transcription factors: key regulators of hematopoiesis. *Exp Hematol*, 23(2):99–107, 1995. ISSN 0301-472X (Print) 0301-472x. 13

J. Nardelli, D. Thiesson, Y. Fujiwara, F. Y. Tsai, and S. H. Orkin. Expression and genetic interaction of transcription factors gata-2 and gata-3 during development of the mouse central nervous system. *Dev Biol*, 210(2):305–21, 1999. ISSN 0012-1606 (Print) 0012-1606. doi: 10.1006/dbio.1999.9278. 13

K. Lillevali, T. Matilainen, A. Karis, and M. Salminen. Partially overlapping expression of gata2 and gata3 during inner ear development. *Dev Dyn*, 231(4):775–81, 2004. ISSN 1058-8388 (Print) 1058-8388. doi: 10.1002/dvdy.20185. 13

J. D. Molkentin. The zinc finger-containing transcription factors gata-4, -5, and -6. ubiquitously expressed regulators of tissue-specific gene expression. *J Biol Chem*, 275(50):38949–52, 2000. ISSN 0021-9258 (Print) 0021-9258. doi: 10.1074/jbc.R000029200. 13

J. D. Molkentin, Q. Lin, S. A. Duncan, and E. N. Olson. Requirement of the transcription factor gata4 for heart tube formation and ventral morphogenesis. *Genes Dev*, 11(8):1061–72, 1997. ISSN 0890-9369 (Print) 0890-9369. doi: 10.1101/gad.11.8.1061. 13

J. D. Molkentin, K. M. Tymitz, J. A. Richardson, and E. N. Olson. Abnormalities of the genitourinary tract in female mice lacking gata5. *Mol Cell Biol*, 20(14):5256–60, 2000. ISSN 0270-7306 (Print) 0270-7306. doi: 10.1128/mcb.20.14.5256-5260.2000. 13

E. E. Morrisey, Z. Tang, K. Sigrist, M. M. Lu, F. Jiang, H. S. Ip, and M. S. Parmacek. Gata6 regulates hnf4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes Dev*, 12(22):3579–90, 1998. ISSN 0890-9369 (Print) 0890-9369. doi: 10.1101/gad.12.22.3579. 13

M. Koutsourakis, A. Langeveld, R. Patient, R. Beddington, and F. Grosveld. The transcription factor gata6 is essential for early extraembryonic development. *Development*, 126(9):723–32, 1999. ISSN 0950-1991 (Print) 0950-1991. 13

K. Decker, D. C. Goldman, C. L. Grasch, and L. Sussel. Gata6 is an important regulator of mouse pancreas development. *Dev Biol*, 298(2):415–29, 2006. ISSN 0012-1606 (Print) 0012-1606. doi: 10.1016/j.ydbio.2006.06.046. 14

I. Ketola, T. Otonkoski, M. A. Pulkkinen, H. Niemi, J. Palgi, C. M. Jacobsen, D. B. Wilson, and M. Heikinheimo. Transcription factor gata-6 is expressed in the endocrine and gata-4 in the exocrine pancreas. *Mol Cell Endocrinol*, 226(1-2):51–7, 2004. ISSN 0303-7207 (Print) 0303-7207. doi: 10.1016/j.mce.2004.06.007. 14

D. J. Sartori, C. J. Wilbur, S. Y. Long, M. M. Rankin, C. Li, J. P. Bradfield, H. Hakonarson, S. F. Grant, W. T. Pu, and J. A. Kushner. Gata factors promote er integrity and beta-cell survival and contribute to type 1 diabetes risk. *Mol Endocrinol*, 28(1):28–39, 2014. ISSN 0888-8809. doi: 10.1210/me.2013-1265. 14

H. L. Allen, S. E. Flanagan, C. Shaw-Smith, E. De Franco, I. Akerman, R. Caswell, J. Ferrer, A. T. Hattersley, and S. Ellard. Gata6 haploinsufficiency causes pancreatic agenesis in humans. *Nat Genet*, 44(1):20–22, 2011. ISSN 1061-4036. doi: 10.1038/ng.1035. 14

C. Y. Chia, P. Madrigal, Slij Denil, I. Martinez, J. Garcia-Bernardo, R. El-Khairi, M. Chhatriwala, M. H. Shepherd, A. T. Hattersley, N. R. Dunn, and L. Vallier. Gata6 cooperates with eomes/smad2/3 to deploy the gene regulatory network governing human definitive endoderm and pancreas formation. *Stem Cell Reports*, 12 (1):57–70, 2019. ISSN 2213-6711. doi: 10.1016/j.stemcr.2018.12.003. 14

A. Tiyaboonchai, F. L. Cardenas-Diaz, L. Ying, J. A. Maguire, X. Sim, C. Jobaliya, A. L. Gagne, S. Kishore, D. E. Stanescu, N. Hughes, D. D. De Leon, D. L. French, and P. Gadue. Gata6 plays an important role in the induction of human definitive endoderm, development of the pancreas, and functionality of pancreatic beta cells. *Stem Cell Reports*, 8(3):589–604, 2017. ISSN 2213-6711. doi: 10.1016/j.stemcr.2016.12.026. 14, 74

M. Carrasco, I. Delgado, B. Soria, F. Martin, and A. Rojas. Gata4 and gata6 control mouse pancreas organogenesis. *J Clin Invest*, 122(10):3504–15, 2012. ISSN 0021-9738. doi: 10.1172/jci63240. 14

K. A. Kwei, M. D. Bashyam, J. Kao, R. Ratheesh, E. C. Reddy, Y. H. Kim, K. Montgomery, C. P. Giacomini, Y. L. Choi, S. Chatterjee, C. A. Karikari, K. Salari, P. Wang, T. Hernandez-Boussard, G. Swarnalata, M. van de Rijn, A. Maitra, and

J. R. Pollack. Genomic profiling identifies gata6 as a candidate oncogene amplified in pancreatobiliary cancer. *PLoS Genet*, 4(5):e1000081, 2008. ISSN 1553-7390. doi: 10.1371/journal.pgen.1000081. 14, 111

B. Fu, M. Luo, S. Lakkur, R. Lucito, and C. A. Iacobuzio-Donahue. Frequent genomic copy number gain and overexpression of gata-6 in pancreatic carcinoma. *Cancer Biol Ther*, 7(10):1593–601, 2008. ISSN 1538-4047. doi: 10.4161/cbt.7.10.6565. 14

Y. Zhong, Z. Wang, B. Fu, F. Pan, S. Yachida, M. Dhara, E. Albesiano, L. Li, Y. Naito, F. Vilardell, C. Cummings, P. Martinelli, A. Li, R. Yonescu, Q. Ma, C. A. Griffin, F. X. Real, and C. A. Iacobuzio-Donahue. Gata6 activates wnt signaling in pancreatic cancer by negatively regulating the wnt antagonist dickkopf-1. *PLoS One*, 6(7):e22129, 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0022129. 14

P. Martinelli, E. Carrillo-de Santa Pau, T. Cox, Jr. Sainz, B., N. Dusetti, W. Greenhalf, L. Rinaldi, E. Costello, P. Ghaneh, N. Malats, M. Buchler, M. Pajic, A. V. Biankin, J. Iovanna, J. Neoptolemos, and F. X. Real. Gata6 regulates emt and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. *Gut*, 66(9):1665–1676, 2017. ISSN 0017-5749. doi: 10.1136/gutjnl-2015-311256. 14

N. Ghosh, U. Hossain, A. Mandal, and P. C. Sil. The wnt signaling pathway: a potential therapeutic target against cancer. *Ann N Y Acad Sci*, 1443(1):54–74, 2019. ISSN 0077-8923. doi: 10.1111/nyas.14027. 15

B. T. MacDonald, K. Tamai, and X. He. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev Cell*, 17(1):9–26, 2009. ISSN 1534-5807. doi: 10.1016/j.devcel.2009.06.016. 15

L. Zhang and J. W. Shay. Multiple roles of apc and its therapeutic implications in colorectal cancer. *J Natl Cancer Inst*, 109(8), 2017. ISSN 0027-8874. doi: 10.1093/jnci/djw332. 15, 134

Y. Zhang, J. P. th Morris, W. Yan, H. K. Schofield, A. Gurney, D. M. Simeone, S. E. Millar, T. Hoey, M. Hebrok, and M. Pasca di Magliano. Canonical wnt signaling is required for pancreatic carcinogenesis. *Cancer Res*, 73(15):4909–22, 2013. ISSN 0008-5472. doi: 10.1158/0008-5472.Can-12-4384. 15

M. Sano, D. R. Driscoll, W. E. DeJesus-Monge, B. Quattrochi, V. A. Appleman, J. Ou, L. J. Zhu, N. Yoshida, S. Yamazaki, T. Takayama, M. Sugitani, N. Nemoto, D. S. Klimstra, and B. C. Lewis. Activation of wnt/beta-catenin signaling enhances pancreatic cancer development and the malignant potential via up-regulation of cyr61. *Neoplasia*, 18(12):785–794, 2016. ISSN 1476-5586. doi: 10.1016/j.neo.2016.11.004. 15

M. Yu, D. T. Ting, S. L. Stott, B. S. Wittner, F. Ozsolak, S. Paul, J. C. Ciciliano, M. E. Smas, D. Winokur, A. J. Gilman, M. J. Ulman, K. Xega, G. Contino, B. Alagesan, B. W. Brannigan, P. M. Milos, D. P. Ryan, L. V. Sequist, N. Bardeesy,

S. Ramaswamy, M. Toner, S. Maheswaran, and D. A. Haber. Rna sequencing of pancreatic circulating tumour cells implicates wnt signalling in metastasis. *Nature*, 487(7408):510–3, 2012. ISSN 0028-0836. doi: 10.1038/nature11217. 15

N. Krishnamurthy and R. Kurzrock. Targeting the wnt/beta-catenin pathway in cancer: Update on effectors and inhibitors. *Cancer Treat Rev*, 62:50–60, 2018. ISSN 0305-7372. doi: 10.1016/j.ctrv.2017.11.002. 15

G. Zeng, M. Germinaro, A. Micsenyi, N. K. Monga, A. Bell, A. Sood, V. Malhotra, N. Sood, V. Midda, D. K. Monga, D. M. Kokkinakis, and S. P. Monga. Aberrant wnt/beta-catenin signaling in pancreatic adenocarcinoma. *Neoplasia*, 8(4):279–89, 2006. ISSN 1476-5586. doi: 10.1593/neo.05607. 15

L. Broutier, A. Andersson-Rolf, C. J. Hindley, S. F. Boj, H. Clevers, B. K. Koo, and M. Huch. Culture and establishment of self-renewing human and mouse adult liver and pancreas 3d organoids and their genetic manipulation. *Nat Protoc*, 11(9): 1724–43, 2016. ISSN 1750-2799. doi: 10.1038/nprot.2016.097. 16

L. M. Murrow, R. J. Weber, and Z. J. Gartner. Dissecting the stem cell niche with organoid models: an engineering-based approach. *Development*, 144(6):998–1007, 2017. ISSN 0950-1991. doi: 10.1242/dev.140905. 16

Waddington C.H. Genetic assimilation of the bithorax phenotype. *International Journal of Organoic Evolution*, X:1–13, 1956. 18

J. B. Gurdon, R. A. Laskey, and O. R. Reeves. The developmental capacity of nuclei transplanted from keratinized skin cells of adult frogs. *J Embryol Exp Morphol*, 34 (1):93–112, 1975. ISSN 0022-0752 (Print) 0022-0752. 18

A. Chatterjee, E. J. Rodger, and M. R. Eccles. Epigenetic drivers of tumourigenesis and cancer metastasis. *Semin Cancer Biol*, 51:149–159, 2018. ISSN 1044-579x. doi: 10.1016/j.semcancer.2017.08.004. 18, 21

O. G. McDonald, X. Li, T. Saunders, R. Tryggvadottir, S. J. Mentch, M. O. Warmoes, A. E. Word, A. Carrer, T. H. Salz, S. Natsume, K. M. Stauffer, A. Makohon-Moore, Y. Zhong, H. Wu, K. E. Wellen, J. W. Locasale, C. A. Iacobuzio-Donahue, and A. P. Feinberg. Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. *Nat Genet*, 49(3):367–376, 2017. ISSN 1061-4036. doi: 10.1038/ng.3753. 18, 50

SK. Zaidi, DW. Young, M. Montecino, JB. Lian, JL. Stein, AJ. van Wijnen, and GS. Stein. Architectural epigenetics: mitotic retention of mammalian transcriptional regulatory information. *Mol Cell Biol*, 30(20), 2010a. ISSN 4758-66. doi: 10.1128/MCB.00646-10. 19

W. Fischle. In nucleo enzymatic assays for the identification and characterization of histone modifying activities. *Methods*, 36(4):362–7, 2005. ISSN 1046-2023 (Print) 1046-2023. doi: 10.1016/j.ymeth.2005.03.008. 20

T. Jenuwein and C. D. Allis. Translating the histone code. *Science*, 293(5532):1074–80, 2001. ISSN 0036-8075 (Print) 0036-8075. doi: 10.1126/science.1063127. 20

X. Dong, M. C. Greven, A. Kundaje, S. Djebali, J. B. Brown, C. Cheng, T. R. Gingeras, M. Gerstein, R. Guigo, E. Birney, and Z. Weng. Modeling gene expression using chromatin features in various cellular contexts. *Genome Biol*, 13(9):R53, 2012. ISSN 1474-7596. doi: 10.1186/gb-2012-13-9-r53. 20

D. Cano-Rodriguez, R. A. Gjaltema, L. J. Jilderda, P. Jellema, J. Dokter-Fokkens, M. H. Ruiters, and M. G. Rots. Writing of h3k4me3 overcomes epigenetic silencing in a sustained but context-dependent manner. *Nat Commun*, 7:12284, 2016. ISSN 2041-1723. doi: 10.1038/ncomms12284. 20, 134

A. P. Feinberg and B. Vogelstein. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, 301(5895):89–92, 1983. ISSN 0028-0836 (Print) 0028-0836. doi: 10.1038/301089a0. 20

S. Saghafinia, M. Mina, N. Riggi, D. Hanahan, and G. Ciriello. Pan-cancer landscape of aberrant dna methylation across human tumors. *Cell Rep*, 25(4):1066–1080.e8, 2018. doi: 10.1016/j.celrep.2018.09.082. 20

T. H. Bestor and V. M. Ingram. Two dna methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with dna. *Proc Natl Acad Sci U S A*, 80(18):5559–63, 1983. ISSN 0027-8424 (Print) 0027-8424. doi: 10.1073/pnas.80.18.5559. 20

M. Curradi, A. Izzo, G. Badaracco, and N. Landsberger. Molecular mechanisms of gene silencing mediated by dna methylation. *Mol Cell Biol*, 22(9):3157–73, 2002. ISSN 0270-7306 (Print) 0270-7306. doi: 10.1128/mcb.22.9.3157-3173.2002. 21

Z. D. Smith and A. Meissner. Dna methylation: roles in mammalian development. *Nat Rev Genet*, 14(3):204–20, 2013. ISSN 1471-0056. doi: 10.1038/nrg3354. 21

D. D. De Carvalho, J. S. You, and P. A. Jones. Dna methylation and cellular reprogramming. *Trends Cell Biol*, 20(10):609–17, 2010. ISSN 0962-8924. doi: 10.1016/j.tcb.2010.08.003. 21

A. Meissner. Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol*, 28(10):1079–88, 2010. ISSN 1087-0156. doi: 10.1038/nbt.1684. 21

J. K. Christman. 5-azacytidine and 5-aza-2'-deoxycytidine as inhibitors of dna methylation: mechanistic studies and their implications for cancer therapy. *Oncogene*, 21(35):5483–95, 2002. ISSN 0950-9232 (Print) 0950-9232. doi: 10.1038/sj.onc.1205699. 21

K. Ghoshal, J. Datta, S. Majumder, S. Bai, H. Kutay, T. Motiwala, and S. T. Jacob. 5-aza-deoxycytidine induces selective degradation of dna methyltransferase 1 by a proteasomal pathway that requires the ken box, bromo-adjacent homology domain,

and nuclear localization signal. *Mol Cell Biol*, 25(11):4727–41, 2005. ISSN 0270-7306 (Print) 0270-7306. doi: 10.1128/mcb.25.11.4727-4741.2005. 21

A. Patra, M. Deb, R. Dahiya, and S. K. Patra. 5-aza-2'-deoxycytidine stress response and apoptosis in prostate cancer. *Clin Epigenetics*, 2(2):339–48, 2011. ISSN 1868-7075. doi: 10.1007/s13148-010-0019-x. 21

P. G. Corn, B. D. Smith, E. S. Ruckdeschel, D. Douglas, S. B. Baylin, and J. G. Herman. E-cadherin expression is silenced by 5' cpg island methylation in acute leukemia. *Clin Cancer Res*, 6(11):4243–8, 2000. ISSN 1078-0432 (Print) 1078-0432. 21

K. Raj and G. J. Mufti. Azacytidine (vidaza(r)) in the treatment of myelodysplastic syndromes. *Ther Clin Risk Manag*, 2(4):377–88, 2006. ISSN 1176-6336 (Print) 1176-6336. doi: 10.2147/tcrm.2006.2.4.377. 21

D. Grasso, J. Bintz, G. Lomberk, M. I. Molejon, C. Loncle, M. N. Garcia, M. B. Lopez, R. Urrutia, and J. L. Iovanna. Pivotal role of the chromatin protein nupr1 in kras-induced senescence and transformation. *Sci Rep*, 5:17549, 2015. ISSN 2045-2322. doi: 10.1038/srep17549. 21

R. Juttermann, E. Li, and R. Jaenisch. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of dna methyltransferase rather than dna demethylation. *Proc Natl Acad Sci U S A*, 91(25):11797–801, 1994. ISSN 0027-8424 (Print) 0027-8424. doi: 10.1073/pnas.91.25.11797. 21

A. S. Yang, K. D. Doshi, S. W. Choi, J. B. Mason, R. K. Mannari, V. Gharybian, R. Luna, A. Rashid, L. Shen, M. R. Estecio, H. M. Kantarjian, G. Garcia-Manero, and J. P. Issa. Dna methylation changes after 5-aza-2'-deoxycytidine therapy in patients with leukemia. *Cancer Res*, 66(10):5495–503, 2006. ISSN 0008-5472 (Print) 0008-5472. doi: 10.1158/0008-5472.Can-05-2385. 21

M. Karahoca and R. L. Momparler. Pharmacokinetic and pharmacodynamic analysis of 5-aza-2'-deoxycytidine (decitabine) in the design of its dose-schedule for cancer therapy. *Clin Epigenetics*, 5(1):3, 2013. ISSN 1868-7075 (Print) 1868-7075. doi: 10.1186/1868-7083-5-3. 21

M. Bostick, J. K. Kim, P. O. Esteve, A. Clark, S. Pradhan, and S. E. Jacobsen. Uhrf1 plays a role in maintaining dna methylation in mammalian cells. *Science*, 317(5845):1760–4, 2007. ISSN 0036-8075. doi: 10.1126/science.1147939. 22

A. Hermann, H. Gowher, and A. Jeltsch. Biochemistry and biology of mammalian dna methyltransferases. *Cell Mol Life Sci*, 61(19-20):2571–87, 2004. ISSN 1420-682X (Print) 1420-682x. doi: 10.1007/s00018-004-4201-1. 22

S. Seisenberger, J. R. Peat, T. A. Hore, F. Santos, W. Dean, and W. Reik. Reprogramming dna methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philos Trans R Soc Lond B Biol Sci*, 368(1609):20110330, 2013. ISSN 0962-8436. doi: 10.1098/rstb.2011.0330. 22

W. Mayer, A. Niveleau, J. Walter, R. Fundele, and T. Haaf. Demethylation of the zygotic paternal genome. *Nature*, 403(6769):501–2, 2000. ISSN 0028-0836 (Print) 0028-0836. doi: 10.1038/35000656. 22

M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, and A. Rao. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian dna by mll partner tet1. *Science*, 324 (5929):930–5, 2009. ISSN 0036-8075. doi: 10.1126/science.1170116. 22, 73

S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, and Y. Zhang. Role of tet proteins in 5mc to 5hmc conversion, es-cell self-renewal and inner cell mass specification. *Nature*, 466(7310):1129–33, 2010. ISSN 0028-0836. doi: 10.1038/nature09303. 22, 74

Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, and G. L. Xu. Tet-mediated formation of 5-carboxylcytosine and its excision by tdg in mammalian dna. *Science*, 333(6047):1303–7, 2011. ISSN 0036-8075. doi: 10.1126/science.1210944. 22

A. Maiti and A. C. Drohat. Thymine dna glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of cpg sites. *J Biol Chem*, 286(41):35334–8, 2011. ISSN 0021-9258. doi: 10.1074/jbc.C111.284620. 22

A. R. Weber, C. Krawczyk, A. B. Robertson, A. Kusnierczyk, C. B. Vagbo, D. Schuermann, A. Klungland, and P. Schar. Biochemical reconstitution of tet1-tdg-ber-dependent active dna demethylation reveals a highly coordinated mechanism. *Nat Commun*, 7:10806, 2016. ISSN 2041-1723. doi: 10.1038/ncomms10806. 22

S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, and Y. Zhang. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 333(6047):1300–3, 2011. ISSN 0036-8075. doi: 10.1126/science.1210597. 22, 73

H. Hashimoto, Y. Liu, A. K. Upadhyay, Y. Chang, S. B. Howerton, P. M. Vertino, X. Zhang, and X. Cheng. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res*, 40(11):4841–9, 2012. ISSN 0305-1048. doi: 10.1093/nar/gks155. 22, 23

L. Hu, J. Lu, J. Cheng, Q. Rao, Z. Li, H. Hou, Z. Lou, L. Zhang, W. Li, W. Gong, M. Liu, C. Sun, X. Yin, J. Li, X. Tan, P. Wang, Y. Wang, D. Fang, Q. Cui, P. Yang, C. He, H. Jiang, C. Luo, and Y. Xu. Structural insight into substrate preference for tet-mediated oxidation. *Nature*, 527(7576):118–22, 2015. ISSN 0028-0836. doi: 10.1038/nature15713. 22

S. Kriaucionis and N. Heintz. The nuclear dna base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. *Science*, 324(5929):929–30, 2009. ISSN 0036-8075. doi: 10.1126/science.1169786. 22

H. Wu, A. C. D'Alessio, S. Ito, Z. Wang, K. Cui, K. Zhao, Y. E. Sun, and Y. Zhang. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev*, 25(7):679–84, 2011. ISSN 0890-9369. doi: 10.1101/gad.2036011. 22

B. Xia, D. Han, X. Lu, Z. Sun, A. Zhou, Q. Yin, H. Zeng, M. Liu, X. Jiang, W. Xie, C. He, and C. Yi. Bisulfite-free, base-resolution analysis of 5-formylcytosine at the genome scale. *Nat Methods*, 12(11):1047–50, 2015. ISSN 1548-7091. doi: 10.1038/nmeth.3569. 22

K. C. Johnson, E. A. Houseman, J. E. King, K. M. von Herrmann, C. E. Fadul, and B. C. Christensen. 5-hydroxymethylcytosine localizes to enhancer elements and is associated with survival in glioblastoma patients. *Nat Commun*, 7:13177, 2016. ISSN 2041-1723. doi: 10.1038/ncomms13177. 23, 52, 73

A. Inoue and Y. Zhang. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science*, 334(6053):194, 2011. ISSN 0036-8075. doi: 10.1126/science.1212483. 23

D. Ji, K. Lin, J. Song, and Y. Wang. Effects of tet-induced oxidation products of 5-methylcytosine on dnmt1- and dnmt3a-mediated cytosine methylation. *Mol Biosyst*, 10(7):1749–52, 2014. ISSN 1742-2051. doi: 10.1039/c4mb00150h. 23

X. Wu and Y. Zhang. Tet-mediated active dna demethylation: mechanism, function and beyond. *Nat Rev Genet*, 18(9):517–534, 2017. ISSN 1471-0056. doi: 10.1038/nrg.2017.33. 23

K. Skvortsova, E. Zotenko, P. L. Luu, C. M. Gould, S. S. Nair, S. J. Clark, and C. Stirzaker. Comprehensive evaluation of genome-wide 5-hydroxymethylcytosine profiling approaches in human dna. *Epigenetics Chromatin*, 10:16, 2017. ISSN 1756-8935. doi: 10.1186/s13072-017-0123-7. 23

J. Li, X. Wu, Y. Zhou, M. Lee, L. Guo, W. Han, W. Mo, W. M. Cao, D. Sun, R. Xie, and Y. Huang. Decoding the dynamic dna methylation and hydroxymethylation landscapes in endodermal lineage intermediates during pancreatic differentiation of hesc. *Nucleic Acids Res*, 46(6):2883–2900, 2018. ISSN 0305-1048. doi: 10.1093/nar/gky063. 24, 25, 62, 63, 64, 74, 76, 81, 129, 130

A. Tovy, A. Spiro, R. McCarthy, Z. Shipony, Y. Aylon, K. Allton, E. Ainbinder, N. Furth, A. Tanay, M. Barton, and M. Oren. p53 is essential for dna methylation homeostasis in naive embryonic stem cells, and its loss promotes clonal heterogeneity. *Genes Dev*, 31(10):959–972, 2017. ISSN 0890-9369. doi: 10.1101/gad.299198.117. 24, 74, 127, 130, 133

W. A. Pastor, L. Aravind, and A. Rao. Tetonic shift: biological roles of tet proteins in dna demethylation and transcription. *Nat Rev Mol Cell Biol*, 14(6):341–56, 2013. ISSN 1471-0072. doi: 10.1038/nrm3589. 25

L. Hu, Z. Li, J. Cheng, Q. Rao, W. Gong, M. Liu, Y. G. Shi, J. Zhu, P. Wang, and Y. Xu. Crystal structure of tet2-dna complex: insight into tet-mediated 5mc oxidation. *Cell*, 155(7):1545–55, 2013. ISSN 0092-8674. doi: 10.1016/j.cell.2013.11.020. 25

H. Zhang, X. Zhang, E. Clark, M. Mulcahey, S. Huang, and Y. G. Shi. Tet1 is a dna-binding protein that modulates dna methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res*, 20(12):1390–3, 2010. ISSN 1001-0602. doi: 10.1038/cr.2010.156. 25

Y. Xu, F. Wu, L. Tan, L. Kong, L. Xiong, J. Deng, A. J. Barbera, L. Zheng, H. Zhang, S. Huang, J. Min, T. Nicholson, T. Chen, G. Xu, Y. Shi, K. Zhang, and Y. G. Shi. Genome-wide regulation of 5hmc, 5mc, and gene expression by tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell*, 42(4):451–64, 2011. ISSN 1097-2765. doi: 10.1016/j.molcel.2011.04.005. 25

K. D. Rasmussen and K. Helin. Role of tet enzymes in dna methylation, development, and cancer. *Genes Dev*, 30(7):733–50, 2016. ISSN 0890-9369. doi: 10.1101/gad.276568.115. 25, 26, 27

Y. Wang, M. Xiao, X. Chen, L. Chen, Y. Xu, L. Lv, P. Wang, H. Yang, S. Ma, H. Lin, B. Jiao, R. Ren, D. Ye, K. L. Guan, and Y. Xiong. Wt1 recruits tet2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol Cell*, 57(4):662–673, 2015. ISSN 1097-2765. doi: 10.1016/j.molcel.2014.12.023. 25

Y. Chu, Z. Zhao, D. W. Sant, G. Zhu, S. M. Greenblatt, L. Liu, J. Wang, Z. Cao, J. C. Tho, S. Chen, X. Liu, P. Zhang, J. P. Maciejewski, S. Nimer, G. Wang, W. Yuan, F. C. Yang, and M. Xu. Tet2 regulates osteoclast differentiation by interacting with runx1 and maintaining genomic 5-hydroxymethylcytosine (5hmc). *Genomics Proteomics Bioinformatics*, 16(3):172–186, 2018. ISSN 1672-0229. doi: 10.1016/j.gpb.2018.04.005. 25

H. Nakajima and H. Kunimoto. Tet2 as an epigenetic master regulator for normal and malignant hematopoiesis. *Cancer Sci*, 9(1), 2014. ISSN 1093-9. doi: 10.1111/cas.12484. 26

R. Ono, T. Taki, T. Taketani, M. Taniwaki, H. Kobayashi, and Y. Hayashi. Lcx, leukemia-associated protein with a cxxc domain, is fused to mll in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res*, 62(14):4075–80, 2002. ISSN 0008-5472 (Print) 0008-5472. 26

R. B. Lorsbach, J. Moore, S. Mathew, S. C. Raimondi, S. T. Mukatira, and J. R. Downing. Tet1, a member of a novel protein family, is fused to mll in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia*, 17(3):637–41, 2003. ISSN 0887-6924 (Print) 0887-6924. doi: 10.1038/sj.leu.2402834. 26

S. Weissmann, T. Alpermann, V. Grossmann, A. Kowarsch, N. Nadarajah, C. Eder, F. Dicker, A. Fasan, C. Haferlach, T. Haferlach, W. Kern, S. Schnittger, and

A. Kohlmann. Landscape of tet2 mutations in acute myeloid leukemia. *Leukemia*, 26(5):934–42, 2012. ISSN 0887-6924. doi: 10.1038/leu.2011.326. 26

L. Scourzic, E. Mouly, and O. A. Bernard. Tet proteins and the control of cytosine demethylation in cancer. *Genome Med*, 7(1):9, 2015. ISSN 1756-994X (Print) 1756-994x. doi: 10.1186/s13073-015-0134-6. 26

M. Ko, Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski, and A. Rao. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant tet2. *Nature*, 468(7325):839–43, 2010. ISSN 0028-0836. doi: 10.1038/nature09586. 26

W. C. Chou, S. C. Chou, C. Y. Liu, C. Y. Chen, H. A. Hou, Y. Y. Kuo, M. C. Lee, B. S. Ko, J. L. Tang, M. Yao, W. Tsay, S. J. Wu, S. Y. Huang, S. C. Hsu, Y. C. Chen, Y. C. Chang, Y. Y. Kuo, K. T. Kuo, F. Y. Lee, M. C. Liu, C. W. Liu, M. H. Tseng, C. F. Huang, and H. F. Tien. Tet2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. *Blood*, 118 (14):3803–10, 2011. ISSN 0006-4971. doi: 10.1182/blood-2011-02-339747. 26

K. Moran-Crusio, L. Reavie, A. Shih, O. Abdel-Wahab, D. Ndiaye-Lobry, C. Lobry, M. E. Figueroa, A. Vasanthakumar, J. Patel, X. Zhao, F. Perna, S. Pandey, J. Madzo, C. Song, Q. Dai, C. He, S. Ibrahim, M. Beran, J. Zavadil, S. D. Nimer, A. Melnick, L. A. Godley, I. Aifantis, and R. L. Levine. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell*, 20 (1):11–24, 2011. ISSN 1535-6108. doi: 10.1016/j.ccr.2011.06.001. 26

M. C. Haffner, A. Chaux, A. K. Meeker, D. M. Esopi, J. Gerber, L. G. Pellakuru, A. Toubaji, P. Argani, C. Iacobuzio-Donahue, W. G. Nelson, G. J. Netto, A. M. De Marzo, and S. Yegnasubramanian. Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget*, 2(8):627–37, 2011. ISSN 1949-2553. doi: 10.18632/oncotarget.316. 27

C. Liu, L. Liu, X. Chen, J. Shen, J. Shan, Y. Xu, Z. Yang, L. Wu, F. Xia, P. Bie, Y. Cui, X. W. Bian, and C. Qian. Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of tet1. *PLoS One*, 8(5):e62828, 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0062828. 27, 93, 135, 136

H. Yang, Y. Liu, F. Bai, J. Y. Zhang, S. H. Ma, J. Liu, Z. D. Xu, H. G. Zhu, Z. Q. Ling, D. Ye, K. L. Guan, and Y. Xiong. Tumor development is associated with decrease of tet gene expression and 5-methylcytosine hydroxylation. *Oncogene*, 32 (5):663–9, 2013. ISSN 0950-9232. doi: 10.1038/onc.2012.67. 27

C. G. Lian, Y. Xu, C. Ceol, F. Wu, A. Larson, K. Dresser, W. Xu, L. Tan, Y. Hu, Q. Zhan, C. W. Lee, D. Hu, B. Q. Lian, S. Kleffel, Y. Yang, J. Neiswender, A. J.

Khorasani, R. Fang, C. Lezcano, L. M. Duncan, R. A. Scolyer, J. F. Thompson, H. Kakavand, Y. Houvras, L. I. Zon, Jr. Mihm, M. C., U. B. Kaiser, T. Schatton, B. A. Woda, G. F. Murphy, and Y. G. Shi. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell*, 150(6):1135–46, 2012. ISSN 0092-8674. doi: 10.1016/j.cell.2012.07.033. 27

Dhawan S Gurlo T, Georgia S. Dna hydroxymethylation regulates beta-cell maturation and expansion. *American Diabetes Association*, 2018. 27

S. Bhattacharyya, K. Pradhan, N. Campbell, J. Mazdo, A. Vasantkumar, S. Maqbool, T. D. Bhagat, S. Gupta, M. Suzuki, Y. Yu, J. M. Greally, U. Steidl, J. Bradner, M. Dawlaty, L. Godley, A. Maitra, and A. Verma. Altered hydroxymethylation is seen at regulatory regions in pancreatic cancer and regulates oncogenic pathways. *Genome Res*, 27(11):1830–1842, 2017. ISSN 1088-9051. doi: 10.1101/gr.222794.117. 27

A. P. Fischer and S. L. Miles. Silencing hif-1alpha induces tet2 expression and augments ascorbic acid induced 5-hydroxymethylation of dna in human metastatic melanoma cells. *Biochem Biophys Res Commun*, 490(2):176–181, 2017. ISSN 0006-291x. doi: 10.1016/j.bbrc.2017.06.017. 27

D. Wu, D. Hu, H. Chen, G. Shi, I. S. Fetahu, F. Wu, K. Rabidou, R. Fang, L. Tan, S. Xu, H. Liu, C. Argueta, L. Zhang, F. Mao, G. Yan, J. Chen, Z. Dong, R. Lv, Y. Xu, M. Wang, Y. Ye, S. Zhang, D. Duquette, S. Geng, C. Yin, C. G. Lian, G. F. Murphy, G. K. Adler, R. Garg, L. Lynch, P. Yang, Y. Li, F. Lan, J. Fan, Y. Shi, and Y. G. Shi. Glucose-regulated phosphorylation of tet2 by ampk reveals a pathway linking diabetes to cancer. *Nature*, 559(7715):637–641, 2018. ISSN 0028-0836. doi: 10.1038/s41586-018-0350-5. 27, 78, 83, 92, 135

P. S. Moore, B. Sipos, S. Orlandini, C. Sorio, F. X. Real, N. R. Lemoine, T. Gress, C. Bassi, G. Kloppel, H. Kalthoff, H. Ungefroren, M. Lohr, and A. Scarpa. Genetic profile of 22 pancreatic carcinoma cell lines. analysis of k-ras, p53, p16 and dpc4/smad4. *Virchows Arch*, 439(6):798–802, 2001. ISSN 0945-6317 (Print) 0945-6317. 30, 119

E. L. Deer, J. Gonzalez-Hernandez, J. D. Coursen, J. E. Shea, J. Ngatia, C. L. Scaife, M. A. Firpo, and S. J. Mulvihill. Phenotype and genotype of pancreatic cancer cell lines. *Pancreas*, 39(4):425–35, 2010. ISSN 0885-3177. doi: 10.1097/MPA.0b013e3181c15963. 30, 119

C. J. David, Y. H. Huang, M. Chen, J. Su, Y. Zou, N. Bardeesy, C. A. Iacobuzio-Donahue, and J. Massague. Tgf-beta tumor suppression through a lethal emt. *Cell*, 164(5):1015–30, 2016. ISSN 0092-8674. doi: 10.1016/j.cell.2016.01.009. 38, 116, 125, 126

T. Webster A. Yang Z. Beck S. Feber A. Teschendorff A. Tian, Y. Morris. Champ: updated methylation analysis pipeline for illumina beadchips. *Bioinformatics*, 33 (24), 2017. ISSN 3982-3984. doi: doi.org/10.1093/bioinformatics/btx513. 38

R. Guinney J Hanzelmann, S. Castelo. Gsva: gene set variation analysis for microarray and rna-seq data. *BMC Bioinformatics*, 14(7), 2013. ISSN 1471-2105. doi: doi:10.1186/1471-2105-14-7. 38

J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehar, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J. E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jane-Valbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I. H. Engels, J. Cheng, G. K. Yu, J. Yu, Jr. Aspesi, P., M. de Silva, K. Jagtap, M. D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R. C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J. P. Mesirov, S. B. Gabriel, G. Getz, K. Ardlie, V. Chan, V. E. Myer, B. L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T. R. Golub, M. P. Morrissey, W. R. Sellers, R. Schlegel, and L. A. Garraway. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, 483(7391):603–7, 2012. ISSN 0028-0836. doi: 10.1038/nature11003. 39, 79

R. Shen, A. B. Olshen, and M. Ladanyi. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics*, 25(22):2906–12, 2009. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp543. 39, 66

SK. Zaidi, DW. Young, M. Montecino, JB. Lian, JL. Stein, AJ. van Wijnen, and GS. Stein. Primerbank: a resource of human and mouse pcr primer pairs for gene expression detection and quantification. *Nucleic acids research*, 38(1), 2010b. ISSN 792-799. doi: doi:10.1093/nar/gkp1005. 44

E. Afgan, D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Cech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Gruning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko, and J. Goecks. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res*, 44(W1): W3–w10, 2016. ISSN 0305-1048. doi: 10.1093/nar/gkw343. 48

M. Lawrence, W. Huber, H. Pages, P. Aboyoun, M. Carlson, R. Gentleman, M. T. Morgan, and V. J. Carey. Software for computing and annotating genomic ranges. *PLoS Comput Biol*, 9(8):e1003118, 2013. ISSN 1553-734x. doi: 10.1371/journal.pcbi.1003118. 48

S. Anders and W. Huber. Differential expression analysis for sequence count data. *Genome Biol*, 11(10):R106, 2010. ISSN 1474-7596. doi: 10.1186/gb-2010-11-10-r106. 48

R. J. C. Kluin, K. Kemper, T. Kuilman, J. R. de Ruiter, V. Iyer, J. V. Forment, P. Cornelissen-Steijger, I. de Rink, P. Ter Brugge, J. Y. Song, S. Klarenbeek, U. McDermott, J. Jonkers, A. Velds, D. J. Adams, D. S. Peeper, and O. Krijgsman. Xenofilter: computational deconvolution of mouse and human reads in tumor

xenograft sequence data. *BMC Bioinformatics*, 19(1):366, 2018. ISSN 1471-2105. doi: 10.1186/s12859-018-2353-5. 48, 104

A. Z. Ayob and T. S. Ramasamy. Cancer stem cells as key drivers of tumour progression. *J Biomed Sci*, 25(1):20, 2018. ISSN 1021-7770. doi: 10.1186/s12929-018-0426-4. 50

M. Frommer, L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy, and C. L. Paul. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual dna strands. *Proc Natl Acad Sci U S A*, 89 (5):1827–31, 1992. ISSN 0027-8424 (Print) 0027-8424. doi: 10.1073/pnas.89.5.1827. 51

Y. Huang, W. A. Pastor, Y. Shen, M. Tahiliani, D. R. Liu, and A. Rao. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One*, 5(1):e8888, 2010. ISSN 1932-6203. doi: 10.1371/journal.pone.0008888. 51

M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik, and S. Balasubramanian. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, 336(6083):934–7, 2012. ISSN 0036-8075. doi: 10.1126/science.1220671. 52

R. Lewis, J. A. Drebin, M. P. Callery, D. Fraker, T. S. Kent, J. Gates, and Jr. Vollmer, C. M. A contemporary analysis of survival for resected pancreatic ductal adenocarcinoma. *HPB (Oxford)*, 15(1):49–60, 2013. ISSN 1365-182x. doi: 10.1111/j.1477-2574.2012.00571.x. 55

P. Robbe, N. Popitsch, S. J. L. Knight, P. Antoniou, J. Becq, M. He, A. Kanapin, A. Samsonova, D. V. Vavoulis, M. T. Ross, Z. Kingsbury, M. Cabes, S. D. C. Ramos, S. Page, H. Dreau, K. Ridout, L. J. Jones, A. Tuff-Lacey, S. Henderson, J. Mason, F. M. Buffa, C. Verrill, D. Maldonado-Perez, I. Roxanis, E. Collantes, L. Browning, S. Dhar, S. Damato, S. Davies, M. Caulfield, D. R. Bentley, J. C. Taylor, C. Turnbull, and A. Schuh. Clinical whole-genome sequencing from routine formalin-fixed, paraffin-embedded specimens: pilot study for the 100,000 genomes project. *Genet Med*, 20(10):1196–1205, 2018. ISSN 1098-3600. doi: 10.1038/gim.2017.241. 57

S. K. Stewart, T. J. Morris, P. Guilhamon, H. Bulstrode, M. Bachman, S. Balasubramanian, and S. Beck. oxbs-450k: a method for analysing hydroxymethylation using 450k beadchips. *Methods*, 72:9–15, 2015. ISSN 1046-2023. doi: 10.1016/j.ymeth.2014.08.009. 61, 62, 73

L. Wiehle, G. Raddatz, T. Musch, M. M. Dawlaty, R. Jaenisch, F. Lyko, and A. Breiling. Tet1 and tet2 protect dna methylation canyons against hypermethylation. *Mol Cell Biol*, 36(3):452–61, 2016. ISSN 0270-7306. doi: 10.1128/mcb.00587-15. 63

M. M. Dawlaty, A. Breiling, T. Le, G. Raddatz, M. I. Barrasa, A. W. Cheng, Q. Gao, B. E. Powell, Z. Li, M. Xu, K. F. Faull, F. Lyko, and R. Jaenisch. Combined

deficiency of tet1 and tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell*, 24(3):310–23, 2013. ISSN 1534-5807. doi: 10.1016/j.devcel.2012.12.015. 63

Clarke A Bird A Berger J, Owen S. Mbd2 is required for correct spatial gene expression in the gut. *Mol Cell Biol*, 27:4049–4057, 2007. 67

M. F. Rodrigues, C. M. Esteves, F. C. Xavier, and F. D. Nunes. Methylation status of homeobox genes in common human cancers. *Genomics*, 108(5-6):185–193, 2016. ISSN 0888-7543. doi: 10.1016/j.ygeno.2016.11.001. 70

S. Gray, H. S. Pandha, A. Michael, G. Middleton, and R. Morgan. Hox genes in pancreatic development and cancer. *Jop*, 12(3):216–9, 2011. ISSN 1590-8577. 70

M. T. Bocker, F. Tuorto, G. Raddatz, T. Musch, F. C. Yang, M. Xu, F. Lyko, and A. Breiling. Hydroxylation of 5-methylcytosine by tet2 maintains the active state of the mammalian hoxa cluster. *Nat Commun*, 3:818, 2012. ISSN 2041-1723. doi: 10.1038/ncomms1826. 70

L. Qiu, F. Liu, S. Yi, X. Li, X. Liu, C. Xiao, C. G. Lian, P. Tu, and Y. Wang. Loss of 5-hydroxymethylcytosine is an epigenetic biomarker in cutaneous t-cell lymphoma. *J Invest Dermatol*, 138(11):2388–2397, 2018. ISSN 0022-202x. doi: 10.1016/j.jid.2018.05.007. 73

A. Zhong, Y. Tian, H. Zhang, and M. Lai. Dna hydroxymethylation of colorectal primary carcinoma and its association with survival. *J Surg Oncol*, 117(5):1029–1037, 2018. ISSN 0022-4790. doi: 10.1002/jso.24978. 73

J. P. Thomson, J. M. Hunter, C. E. Nestor, D. S. Dunican, R. Terranova, J. G. Moggs, and R. R. Meehan. Comparative analysis of affinity-based 5-hydroxymethylation enrichment techniques. *Nucleic Acids Res*, 41(22):e206, 2013. ISSN 0305-1048. doi: 10.1093/nar/gkt1080. 73

D. Globisch, M. Munzel, M. Muller, S. Michalakis, M. Wagner, S. Koch, T. Bruckl, M. Biel, and T. Carell. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One*, 5(12):e15367, 2010. ISSN 1932-6203. doi: 10.1371/journal.pone.0015367. 73

S. Xuan and L. Sussel. Gata4 and gata6 regulate pancreatic endoderm identity through inhibition of hedgehog signaling. *Development*, 143(5):780–6, 2016. ISSN 0950-1991. doi: 10.1242/dev.127217. 74

C. Kuiper and M. C. Vissers. Ascorbate as a co-factor for fe- and 2-oxoglutarate dependent dioxygenases: physiological activity in tumor growth and progression. *Front Oncol*, 4:359, 2014. ISSN 2234-943X (Print) 2234-943x. doi: 10.3389/fonc.2014.00359. 76, 137

T. A. Hore. Modulating epigenetic memory through vitamins and tet: implications for regenerative medicine and cancer treatment. *Epigenomics*, 9(6):863–871, 2017. ISSN 1750-192x. doi: 10.2217/epi-2017-0021. 77

R. Yin, S. Q. Mao, B. Zhao, Z. Chong, Y. Yang, C. Zhao, D. Zhang, H. Huang, J. Gao, Z. Li, Y. Jiao, C. Li, S. Liu, D. Wu, W. Gu, Y. G. Yang, G. L. Xu, and H. Wang. Ascorbic acid enhances tet-mediated 5-methylcytosine oxidation and promotes dna demethylation in mammals. *J Am Chem Soc*, 135(28):10396–403, 2013. ISSN 0002-7863. doi: 10.1021/ja4028346. 77

T. A. Hore, F. von Meyenn, M. Ravichandran, M. Bachman, G. Ficz, D. Oxley, F. Santos, S. Balasubramanian, T. P. Jurkowski, and W. Reik. Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naive pluripotency by complementary mechanisms. *Proc Natl Acad Sci U S A*, 113(43): 12202–12207, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1608679113. 77

H. Wu and Y. Zhang. Reversing dna methylation: mechanisms, genomics, and biological functions. *Cell*, 156(1-2):45–68, 2014. ISSN 0092-8674. doi: 10.1016/j.cell.2013.12.019. 77

Y. Liu, S. Xu, T. Zu, F. Li, S. Sang, C. Liu, Y. An, B. Mi, D. P. Orgill, G. F. Murphy, and C. G. Lian. Reversal of tet-mediated 5-hmc loss in hypoxic fibroblasts by ascorbic acid. *Lab Invest*, 2019. ISSN 0023-6837. doi: 10.1038/s41374-019-0235-8. 77

Z. W. Zhang, M. Abdullahi, and M. J. Farthing. Effect of physiological concentrations of vitamin c on gastric cancer cells and helicobacter pylori. *Gut*, 50(2):165–9, 2002. ISSN 0017-5749 (Print) 0017-5749. 77

Q. Chen, M. G. Espey, A. Y. Sun, J. H. Lee, M. C. Krishna, E. Shacter, P. L. Choyke, C. Pooput, K. L. Kirk, G. R. Buettner, and M. Levine. Ascorbate in pharmacologic concentrations selectively generates ascorbate radical and hydrogen peroxide in extracellular fluid in vivo. *Proc Natl Acad Sci U S A*, 104(21):8749–54, 2007. ISSN 0027-8424 (Print) 0027-8424. doi: 10.1073/pnas.0702854104. 77

C. D'Aniello, F. Cermola, E. J. Patriarca, and G. Minchiotti. Vitamin c in stem cell biology: Impact on extracellular matrix homeostasis and epigenetics. *Stem Cells Int*, 2017:8936156, 2017. ISSN 1687-966X (Print). doi: 10.1155/2017/8936156. 77

K. Polireddy, R. Dong, G. Reed, J. Yu, P. Chen, S. Williamson, P. C. Violet, Z. Pessetto, A. K. Godwin, F. Fan, M. Levine, J. A. Drisko, and Q. Chen. High dose parenteral ascorbate inhibited pancreatic cancer growth and metastasis: Mechanisms and a phase i/iia study. *Sci Rep*, 7(1):17188, 2017. ISSN 2045-2322. doi: 10.1038/s41598-017-17568-8. 77

Q. Chen, M. G. Espey, M. C. Krishna, J. B. Mitchell, C. P. Corpe, G. R. Buettner, E. Shacter, and M. Levine. Pharmacologic ascorbic acid concentrations selectively kill cancer cells: action as a pro-drug to deliver hydrogen peroxide to tissues. *Proc*

Natl Acad Sci U S A, 102(38):13604–9, 2005. ISSN 0027-8424 (Print) 0027-8424. doi: 10.1073/pnas.0506390102. 77

J. A. Drisko, O. K. Serrano, L. R. Spruce, Q. Chen, and M. Levine. Treatment of pancreatic cancer with intravenous vitamin c: a case report. *Anticancer Drugs*, 29 (4):373–379, 2018. ISSN 0959-4973. doi: 10.1097/cad.0000000000000603. 77

Miranda-Massar J Lopez D Duconge J Rodriguez J Adrover P Gonzalez M, Berdiel M. High dose intravenous vitamin c and metastatic pancreatic cancer: Two cases. *Integr Cancer Sci Therap*, 3(10), 2016. 77

J. L. Welsh, B. A. Wagner, T. J. van't Erve, P. S. Zehr, D. J. Berg, T. R. Halfdanarson, N. S. Yee, K. L. Bodeker, J. Du, 2nd Roberts, L. J., J. Drisko, M. Levine, G. R. Buettner, and J. J. Cullen. Pharmacological ascorbate with gemcitabine for the control of metastatic and node-positive pancreatic cancer (pacman): results from a phase i clinical trial. *Cancer Chemother Pharmacol*, 71(3):765–75, 2013. ISSN 0344-5704. doi: 10.1007/s00280-013-2070-8. 77

D. G. Hardie. Amp-activated protein kinase: an energy sensor that regulates all aspects of cell function. *Genes Dev*, 25(18):1895–908, 2011. ISSN 0890-9369. doi: 10.1101/gad.17420111. 78

K. Chen, W. Qian, J. Li, Z. Jiang, L. Cheng, B. Yan, J. Cao, L. Sun, C. Zhou, M. Lei, W. Duan, J. Ma, Q. Ma, and Z. Ma. Loss of ampk activation promotes the invasion and metastasis of pancreatic cancer through an hsf1-dependent pathway. *Mol Oncol*, 11(10):1475–1492, 2017. ISSN 1574-7891. doi: 10.1002/1878-0261.12116. 78, 134

J. Zhu, Y. Chen, Y. Ji, Y. Yu, Y. Jin, X. Zhang, and J. Zhou. Gemcitabine induces apoptosis and autophagy via the ampk/mtor signaling pathway in pancreatic cancer cells. *Biotechnol Appl Biochem*, 65(5):665–671, 2018. ISSN 0885-4513. doi: 10.1002/bab.1657. 78

K. A. Coughlan, R. J. Valentine, N. B. Ruderman, and A. K. Saha. Ampk activation: a therapeutic target for type 2 diabetes? *Diabetes Metab Syndr Obes*, 7:241–53, 2014. ISSN 1178-7007 (Print) 1178-7007. doi: 10.2147/dmso.S43731. 78

J. Permert, I. Ihse, L. Jorfeldt, H. von Schenck, H. J. Arnqvist, and J. Larsson. Pancreatic cancer is associated with impaired glucose metabolism. *Eur J Surg*, 159 (2):101–7, 1993. ISSN 1102-4151 (Print) 1102-4151. 78, 135, 138

S. S. Schwarts, A. Zeidler, A. R. Moossa, S. F. Kuku, and A. H. Rubenstein. A prospective study of glucose tolerance, insulin, c-peptide, and glucagon responses in patients with pancreatic carcinoma. *Am J Dig Dis*, 23(12):1107–14, 1978. ISSN 0002-9211 (Print) 0002-9211. 78

S. F. Mekan, M. M. Safa, R. S. Komrokji, Z. A. Nahleh, and A. R. Jazieh. Is hyperglycemia a prognostic factor in pancreatic cancer? *Journal of Clinical Oncology*, 24 $(18_{s}uppl) : 14078 - -14078, 2006. doi : .$ URL `https://ascopubs.org/doi/abs/10.1200/jco.2006.24.18_suppl.14078`. 78, 92, 135

Z. Wang, S. T. Lai, L. Xie, J. D. Zhao, N. Y. Ma, J. Zhu, Z. G. Ren, and G. L. Jiang. Metformin is associated with reduced risk of pancreatic cancer in patients with type 2 diabetes mellitus: a systematic review and meta-analysis. *Diabetes Res Clin Pract*, 106(1):19–26, 2014. ISSN 0168-8227. 10.1016/j.diabres.2014.04.007. 78

N. Sadeghi, J. L. Abbruzzese, S. C. Yeung, M. Hassan, and D. Li. Metformin use is associated with better survival of diabetic patients with pancreatic cancer. *Clin Cancer Res*, 18(10):2905–12, 2012. ISSN 1078-0432 (Print) 1078-0432. 10.1158/1078-0432.Ccr-11-2994. 78

J. M. Evans, L. A. Donnelly, A. M. Emslie-Smith, D. R. Alessi, and A. D. Morris. Metformin and reduced risk of cancer in diabetic patients. *Bmj*, 330(7503):1304–5, 2005. ISSN 0959-8138. 10.1136/bmj.38415.708634.F7. 78

M. Y. El-Mir, V. Nogueira, E. Fontaine, N. Averet, M. Rigoulet, and X. Leverve. Dimethylbiguanide inhibits cell respiration via an indirect effect targeted on the respiratory chain complex i. *J Biol Chem*, 275(1):223–8, 2000. ISSN 0021-9258 (Print) 0021-9258. 78

J. Andricovich, S. Perkail, Y. Kai, N. Casasanta, W. Peng, and A. Tzatsos. Loss of kdm6a activates super-enhancers to induce gender-specific squamous-like pancreatic cancer and confers sensitivity to bet inhibitors. *Cancer Cell*, 33(3):512–526.e8, 2018. ISSN 1535-6108. 10.1016/j.ccell.2018.02.003. 82

R. J. Dowling, S. Niraula, V. Stambolic, and P. J. Goodwin. Metformin in cancer: translational challenges. *J Mol Endocrinol*, 48(3):R31–43, 2012. ISSN 0952-5041. 10.1530/jme-12-0007. 83

F. Barbieri, S. Thellung, A. Ratto, E. Carra, V. Marini, C. Fucile, A. Bajetto, A. Pattarozzi, R. Wurth, M. Gatti, C. Campanella, G. Vito, F. Mattioli, A. Pagano, A. Daga, A. Ferrari, and T. Florio. In vitro and in vivo antiproliferative activity of metformin on stem-like cells isolated from spontaneous canine mammary carcinomas: translational implications for human tumors. *BMC Cancer*, 15:228, 2015. ISSN 1471-2407. 10.1186/s12885-015-1235-8. 83

G. Rena, D. G. Hardie, and E. R. Pearson. The mechanisms of action of metformin. *Diabetologia*, 60(9):1577–1585, 2017. ISSN 0012-186x. 10.1007/s00125-017-4342-z. 83, 93, 106

G. Zhou, R. Myers, Y. Li, Y. Chen, X. Shen, J. Fenyk-Melody, M. Wu, J. Ventre, T. Doebber, N. Fujii, N. Musi, M. F. Hirshman, L. J. Goodyear, and D. E. Moller. Role of amp-activated protein kinase in mechanism of metformin action. *J Clin Invest*, 108(8):1167–74, 2001. ISSN 0021-9738 (Print) 0021-9738. 10.1172/jci13505. 84

L. Cimmino, I. Dolgalev, Y. Wang, A. Yoshimi, G. H. Martin, J. Wang, V. Ng, B. Xia, M. T. Witkowski, M. Mitchell-Flack, I. Grillo, S. Bakogianni, D. Ndiaye-Lobry, M. T. Martin, M. Guillamot, R. S. Banh, M. Xu, M. E. Figueroa, R. A. Dickins, O. Abdel-Wahab, C. Y. Park, A. Tsirigos, B. G. Neel, and I. Aifantis. Restoration of tet2 function blocks aberrant self-renewal and leukemia progression. *Cell*, 170(6):1079–1095.e20, 2017. ISSN 0092-8674. 10.1016/j.cell.2017.07.032. 85

H. Q. Duong, J. S. Hwang, H. J. Kim, Y. S. Seong, and I. Bae. Bml-275, an ampk inhibitor, induces dna damage, g2/m arrest and apoptosis in human pancreatic cancer cells. *Int J Oncol*, 41(6):2227–36, 2012. ISSN 1019-6439. 10.3892/ijo.2012.1672. 89

W. C. Liao, Y. K. Tu, M. S. Wu, J. T. Lin, H. P. Wang, and K. L. Chien. Blood glucose concentration and risk of pancreatic cancer: systematic review and dose-response meta-analysis. *Bmj*, 350:g7371, 2015. ISSN 0959-8138. 10.1136/bmj.g7371. 92

M. C. M. Vissers and A. B. Das. Potential mechanisms of action for vitamin c in cancer: Reviewing the evidence. *Front Physiol*, 9:809, 2018. ISSN 1664-042X (Print) 1664-042x. 10.3389/fphys.2018.00809. 93

S. Sengupta, T. R. Peterson, and D. M. Sabatini. Regulation of the mtor complex 1 pathway by nutrients, growth factors, and stress. *Mol Cell*, 40(2):310–22, 2010. ISSN 1097-2765. 10.1016/j.molcel.2010.09.026. 93

M. Yang, S. Teng, C. Ma, Y. Yu, P. Wang, and C. Yi. Ascorbic acid inhibits senescence in mesenchymal stem cells through ros and akt/mtor signaling. *Cytotechnology*, 70(5):1301–1313, 2018. ISSN 0920-9069 (Print) 0920-9069. 10.1007/s10616-018-0220-x. 93

M. Moretti, J. Budni, A. E. Freitas, P. B. Rosa, and A. L. Rodrigues. Antidepressant-like effect of ascorbic acid is associated with the modulation of mammalian target of rapamycin pathway. *J Psychiatr Res*, 48(1):16–24, 2014. ISSN 0022-3956. 10.1016/j.jpsychires.2013.10.014. 93

Y. Xie, Y. Jin, B. L. Merenick, M. Ding, K. M. Fetalvero, R. J. Wagner, A. Mai, S. Gleim, D. F. Tucker, M. J. Birnbaum, B. A. Ballif, A. K. Luciano, W. C. Sessa, E. M. Rzucidlo, R. J. Powell, L. Hou, H. Zhao, J. Hwa, J. Yu, and K. A. Martin. Phosphorylation of gata-6 is required for vascular smooth muscle cell differentiation after mtorc1 inhibition. *Sci Signal*, 8(376):ra44, 2015. ISSN 1945-0877. 10.1126/scisignal.2005482. 93

Y. Lai, X. Wei, S. Lin, L. Qin, L. Cheng, and P. Li. Current status and perspectives of patient-derived xenograft models in cancer research. *J Hematol Oncol*, 10(1):106, 2017. ISSN 1756-8722. 10.1186/s13045-017-0470-7. 95

L. Chen, X. Pan, Y. H. Zhang, X. Hu, K. Feng, T. Huang, and Y. D. Cai. Primary tumor site specificity is preserved in patient-derived tumor xenograft models. *Front Genet*, 10:738, 2019. ISSN 1664-8021 (Print) 1664-8021. 10.3389/fgene.2019.00738. 95

M. P. Kim, D. B. Evans, H. Wang, J. L. Abbruzzese, J. B. Fleming, and G. E. Gallick. Generation of orthotopic and heterotopic human pancreatic cancer xenografts in immunodeficient mice. *Nat Protoc*, 4(11):1670–80, 2009. ISSN 1750-2799. 10.1038/nprot.2009.171. 97

Pamela Moll, Michael Ante, Alexander Seitz, and Torsten Reda. Quantseq 3 mrna sequencing for rna quantification. *Nature Methods*, 11:972, 2014. 10.1038/nmeth.f.376. URL `https://doi.org/10.1038/nmeth.f.376`. 104

R. C. Friedman, K. K. Farh, C. B. Burge, and D. P. Bartel. Most mammalian mrnas are conserved targets of micrornas. *Genome Res*, 19(1):92–105, 2009. ISSN 1088-9051 (Print) 1088-9051. 10.1101/gr.082701.108. 104

H. Thorvaldsdottir, J. T. Robinson, and J. P. Mesirov. Integrative genomics viewer (igv): high-performance genomics data visualization and exploration. *Brief Bioinform*, 14(2):178–92, 2013. ISSN 1467-5463. 10.1093/bib/bbs017. 104

X. Liu, Q. Shen, T. Yu, H. Huang, Z. Zhang, J. Ding, Y. Tang, N. Xu, and S. Yue. Small gtpase arl6 controls rh30 rhabdomyosarcoma cell growth through ciliogenesis and hedgehog signaling. *Cell Biosci*, 6:61, 2016. ISSN 2045-3701 (Print) 2045-3701. 10.1186/s13578-016-0126-2. 106

S. C. Goetz, P. J. Ocbina, and K. V. Anderson. The primary cilium as a hedgehog signal transduction machine. *Methods Cell Biol*, 94:199–222, 2009. ISSN 0091-679X (Print) 0091-679x. 10.1016/s0091-679x(08)94010-3. 106

M. D. Arensman, A. N. Kovochich, R. M. Kulikauskas, A. R. Lay, P. T. Yang, X. Li, T. Donahue, M. B. Major, R. T. Moon, A. J. Chien, and D. W. Dawson. Wnt7b mediates autocrine wnt/beta-catenin signaling and anchorage-independent growth in pancreatic adenocarcinoma. *Oncogene*, 33(7):899–908, 2014. ISSN 0950-9232. 10.1038/onc.2013.23. 107

K. Kretzschmar and H. Clevers. Wnt/beta-catenin signaling in adult mammalian epithelial stem cells. *Dev Biol*, 428(2):273–282, 2017. ISSN 0012-1606. 10.1016/j.ydbio.2017.05.015. 109

S. Bruno and Z. Darzynkiewicz. Cell cycle dependent expression and stability of the nuclear protein detected by ki-67 antibody in hl-60 cells. *Cell Prolif*, 25(1):31–40, 1992. ISSN 0960-7722 (Print) 0960-7722. 10.1111/j.1365-2184.1992.tb01435.x. 111

K. S. Inman, A. A. Francis, and N. R. Murray. Complex role for the immune system in initiation and progression of pancreatic cancer. *World J Gastroenterol*, 20(32): 11160–81, 2014. ISSN 1007-9327. 10.3748/wjg.v20.i32.11160. 113

S. A. Hahn, M. Schutte, A. T. Hoque, C. A. Moskaluk, L. T. da Costa, E. Rozenblum, C. L. Weinstein, A. Fischer, C. J. Yeo, R. H. Hruban, and S. E. Kern. Dpc4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science*, 271(5247): 350–3, 1996. ISSN 0036-8075 (Print) 0036-8075. 10.1126/science.271.5247.350. 115

A. Blackford, O. K. Serrano, C. L. Wolfgang, G. Parmigiani, S. Jones, X. Zhang, D. W. Parsons, J. C. Lin, R. J. Leary, J. R. Eshleman, M. Goggins, E. M. Jaffee, C. A. Iacobuzio-Donahue, A. Maitra, J. L. Cameron, K. Olino, R. Schulick, J. Winter, J. M. Herman, D. Laheru, A. P. Klein, B. Vogelstein, K. W. Kinzler, V. E. Velculescu, and R. H. Hruban. Smad4 gene mutations are associated with poor prognosis in pancreatic cancer. *Clin Cancer Res*, 15(14):4674–9, 2009. ISSN 1078-0432 (Print) 1078-0432. 10.1158/1078-0432.Ccr-09-0227. 115

X. Shugang, Y. Hongfa, L. Jianpeng, Z. Xu, F. Jingqi, L. Xiangxiang, and L. Wei. Prognostic value of smad4 in pancreatic cancer: A meta-analysis. *Transl Oncol*, 9(1): 1–7, 2016. ISSN 1936-5233 (Print) 1936-5233. 10.1016/j.tranon.2015.11.007. 116

S. Ahmed, A. D. Bradshaw, S. Gera, M. Z. Dewan, and R. Xu. The tgf-beta/smad4 signaling pathway in pancreatic carcinogenesis and its clinical significance. *J Clin Med*, 6(1), 2017. ISSN 2077-0383 (Print) 2077-0383. 10.3390/jcm6010005. 116

M. Zhao, L. Mishra, and C. X. Deng. The role of tgf-beta/smad4 signaling in cancer. *Int J Biol Sci*, 14(2):111–123, 2018. ISSN 1449-2288. 10.7150/ijbs.23230. 116

Z. Yan, S. Winawer, and E. Friedman. Two different signal transduction pathways can be activated by transforming growth factor beta 1 in epithelial cells. *J Biol Chem*, 269(18):13231–7, 1994. ISSN 0021-9258 (Print) 0021-9258. 116

K. M. Mulder and S. L. Morris. Activation of p21ras by transforming growth factor beta in epithelial cells. *J Biol Chem*, 267(8):5029–31, 1992. ISSN 0021-9258 (Print) 0021-9258. 116

J. J. Lebrun. The dual role of tgfbeta in human cancer: From tumor suppression to cancer metastasis. *ISRN Mol Biol*, 2012:381428, 2012. ISSN 2090-7907 (Print) 2090-7907. 10.5402/2012/381428. 116

B. A. Kennedy, D. E. Deatherage, F. Gu, B. Tang, M. W. Chan, K. P. Nephew, T. H. Huang, and V. X. Jin. Chip-seq defined genome-wide map of tgfbeta/smad4 targets: implications with clinical outcome of ovarian cancer. *PLoS One*, 6(7):e22606, 2011. ISSN 1932-6203. 10.1371/journal.pone.0022606. 121, 122

C. A. Davis, B. C. Hitz, C. A. Sloan, E. T. Chan, J. M. Davidson, I. Gabdank, J. A. Hilton, K. Jain, U. K. Baymuradov, A. K. Narayanan, K. C. Onate, K. Graham, S. R. Miyasato, T. R. Dreszer, J. S. Strattan, O. Jolanki, F. Y. Tanaka, and J. M. Cherry. The encyclopedia of dna elements (encode): data portal update. *Nucleic Acids Res*, 46(D1):D794–d801, 2018. ISSN 0305-1048. 10.1093/nar/gkx1081. 121

N. D. Heintzman, R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*, 39(3):311–8, 2007. ISSN 1061-4036 (Print) 1061-4036. 10.1038/ng1966. 121

M. P. Creyghton, A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, and R. Jaenisch. Histone h3k27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*, 107(50):21931–6, 2010. ISSN 0027-8424. 10.1073/pnas.1016071107. 121

B. K. Wu and C. Brenner. Suppression of tet1-dependent dna demethylation is essential for kras-mediated transformation. *Cell Rep*, 9(5):1827–1840, 2014. 10.1016/j.celrep.2014.10.063. 127

E. de Wit and W. de Laat. A decade of 3c technologies: insights into nuclear organization. *Genes Dev*, 26(1):11–24, 2012. ISSN 0890-9369. 10.1101/gad.179804.111. 127

N. Muller, A. Reinacher-Schick, S. Baldus, J. van Hengel, G. Berx, A. Baar, F. van Roy, W. Schmiegel, and I. Schwarte-Waldhoff. Smad4 induces the tumor suppressor

e-cadherin and p-cadherin in colon carcinoma cells. *Oncogene*, 21(39):6049–58, 2002. ISSN 0950-9232 (Print) 0950-9232. 10.1038/sj.onc.1205766. 128

A. M. Waters and C. J. Der. Kras: The critical driver and therapeutic target for pancreatic cancer. *Cold Spring Harb Perspect Med*, 8(9), 2018. ISSN 2157-1422. 10.1101/cshperspect.a031435. 129

S. K. Patra. Ras regulation of dna-methylation and cancer. *Exp Cell Res*, 314(6): 1193–201, 2008. ISSN 0014-4827 (Print) 0014-4827. 10.1016/j.yexcr.2008.01.012. 129

C. Gazin, N. Wajapeyee, S. Gobeil, C. M. Virbasius, and M. R. Green. An elaborate pathway required for ras-mediated epigenetic silencing. *Nature*, 449(7165):1073–7, 2007. ISSN 0028-0836. 10.1038/nature06251. 129

T. J. Grant, K. Hua, and A. Singh. Molecular pathogenesis of pancreatic cancer. *Prog Mol Biol Transl Sci*, 144:241–275, 2016. ISSN 1877-1173. 10.1016/bs.pmbts.2016.09.008. 133

C. Tiffon. The impact of nutrition and environmental epigenetics on human health and disease. *Int J Mol Sci*, 19(11), 2018. ISSN 1422-0067. 10.3390/ijms19113425. 134

M. Romani, M. P. Pistillo, and B. Banelli. Environmental epigenetics: Crossroad between public health, lifestyle, and cancer prevention. *Biomed Res Int*, 2015:587983, 2015. 10.1155/2015/587983. 134

V. V. Ogryzko, R. L. Schiltz, V. Russanova, B. H. Howard, and Y. Nakatani. The transcriptional coactivators p300 and cbp are histone acetyltransferases. *Cell*, 87(5): 953–9, 1996. ISSN 0092-8674 (Print) 0092-8674. 10.1016/s0092-8674(00)82001-2. 134

P. Tropberger, S. Pott, C. Keller, K. Kamieniarz-Gdula, M. Caron, F. Richter, G. Li, G. Mittler, E. T. Liu, M. Buhler, R. Margueron, and R. Schneider. Regulation of transcription through acetylation of h3k122 on the lateral surface of the histone octamer. *Cell*, 152(4):859–72, 2013. ISSN 0092-8674. 10.1016/j.cell.2013.01.032. 134

A. Ito, C. H. Lai, X. Zhao, S. Saito, M. H. Hamilton, E. Appella, and T. P. Yao. p300/cbp-mediated p53 acetylation is commonly induced by p53-activating agents and inhibited by mdm2. *Embo j*, 20(6):1331–40, 2001. ISSN 0261-4189 (Print) 0261-4189. 10.1093/emboj/20.6.1331. 134

J. Luo, F. Su, D. Chen, A. Shiloh, and W. Gu. Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature*, 408(6810):377–81, 2000. ISSN 0028-0836 (Print) 0028-0836. 10.1038/35042612. 134

Y. Li and E. Seto. Hdacs and hdac inhibitors in cancer development and therapy. *Cold Spring Harb Perspect Med*, 6(10), 2016. ISSN 2157-1422. 10.1101/cshperspect.a026831. 134

I. Koutsounas, C. Giaginis, and S. Theocharis. Histone deacetylase inhibitors and pancreatic cancer: are there any promising clinical trials? *World J Gastroenterol*, 19 (8):1173–81, 2013. ISSN 1007-9327. 10.3748/wjg.v19.i8.1173. 134

E. Chan, L. R. Arlinghaus, D. B. Cardin, L. Goff, J. D. Berlin, A. Parikh, R. G. Abramson, T. E. Yankeelov, S. Hiebert, N. Merchant, S. Bhaskara, and A. B. Chakravarthy. Phase i trial of vorinostat added to chemoradiation with capecitabine in pancreatic cancer. *Radiother Oncol*, 119(2):312–8, 2016. ISSN 0167-8140. 10.1016/j.radonc.2016.04.013. 134

R. P. Sah, A. Sharma, S. Nagpal, S. H. Patlolla, A. Sharma, H. Kandlakunta, V. Anani, R. S. Angom, A. K. Kamboj, N. Ahmed, S. Mohapatra, S. Vivekanandhan, K. A. Philbrick, A. Weston, N. Takahashi, J. Kirkland, N. Javeed, A. Matveyenko, M. J. Levy, D. Mukhopadhyay, and S. T. Chari. Phases of metabolic and soft tissue changes in months preceding a diagnosis of pancreatic ductal adenocarcinoma. *Gastroenterology*, 156(6):1742–1752, 2019. ISSN 0016-5085. 10.1053/j.gastro.2019.01.039. 135

M. S. Yoon. The role of mammalian target of rapamycin (mtor) in insulin signaling. *Nutrients*, 9(11), 2017. ISSN 2072-6643. 10.3390/nu9111176. 135

X. Li, T. Li, Z. Liu, S. Gou, and C. Wang. The effect of metformin on survival of patients with pancreatic cancer: a meta-analysis. *Sci Rep*, 7(1):5825, 2017. ISSN 2045-2322. 10.1038/s41598-017-06207-x. 135

S. Iriana, S. Ahmed, J. Gong, A. A. Annamalai, R. Tuli, and A. E. Hendifar. Targeting mtor in pancreatic ductal adenocarcinoma. *Front Oncol*, 6:99, 2016. ISSN 2234-943X (Print) 2234-943x. 10.3389/fonc.2016.00099. 136

A. Laitala and J. T. Erler. Hypoxic signalling in tumour stroma. *Front Oncol*, 8: 189, 2018. ISSN 2234-943X (Print) 2234-943x. 10.3389/fonc.2018.00189. 136

O. McGinn, V. K. Gupta, P. Dauer, N. Arora, N. Sharma, A. Nomura, V. Dudeja, A. Saluja, and S. Banerjee. Inhibition of hypoxic response decreases stemness and reduces tumorigenic signaling due to impaired assembly of hif1 transcription complex in pancreatic cancer. *Sci Rep*, 7(1):7872, 2017. ISSN 2045-2322. 10.1038/s41598-017-08447-3. 136

P. Jaakkola, D. R. Mole, Y. M. Tian, M. I. Wilson, J. Gielbert, S. J. Gaskell, A. von Kriegsheim, H. F. Hebestreit, M. Mukherji, C. J. Schofield, P. H. Maxwell, C. W. Pugh, and P. J. Ratcliffe. Targeting of hif-alpha to the von hippel-lindau ubiquitylation complex by o2-regulated prolyl hydroxylation. *Science*, 292(5516):468–72, 2001. ISSN 0036-8075 (Print) 0036-8075. 10.1126/science.1059796. 136

P. C. Mahon, K. Hirota, and G. L. Semenza. Fih-1: a novel protein that interacts with hif-1alpha and vhl to mediate repression of hif-1 transcriptional activity. *Genes Dev*, 15(20):2675–86, 2001. ISSN 0890-9369 (Print) 0890-9369. 10.1101/gad.924501. 136

B. Thienpont, E. Galle, and D. Lambrechts. Tet enzymes as oxygen-dependent tumor suppressors: exciting new avenues for cancer management. *Epigenomics*, 8 (11):1445–1448, 2016. ISSN 1750-192x. 10.2217/epi-2016-0126. 137

S. Burr, A. Caldwell, M. Chong, M. Beretta, S. Metcalf, M. Hancock, M. Arno, S. Balu, V. L. Kropf, R. K. Mistry, A. M. Shah, G. E. Mann, and A. C. Brewer. Oxygen gradients can determine epigenetic asymmetry and cellular differentiation via differential regulation of tet activity in embryonic stem cells. *Nucleic Acids Res*, 46(3):1210–1226, 2018. ISSN 0305-1048. 10.1093/nar/gkx1197. 137

S. L. Miles, A. P. Fischer, S. J. Joshi, and R. M. Niles. Ascorbic acid and ascorbate-2-phosphate decrease hif activity and malignant properties of human melanoma cells. *BMC Cancer*, 15:867, 2015. ISSN 1471-2407. 10.1186/s12885-015-1878-5. 137

X. Zhou, J. Chen, G. Yi, M. Deng, H. Liu, M. Liang, B. Shi, X. Fu, Y. Chen, L. Chen, Z. He, J. Wang, and J. Liu. Metformin suppresses hypoxia-induced stabilization of hif-1alpha through reprogramming of oxygen metabolism in hepatocellular carcinoma. *Oncotarget*, 7(1):873–84, 2016. ISSN 1949-2553. 10.18632/oncotarget.6418. 137

H. Fujita, K. Hirose, M. Sato, I. Fujioka, T. Fujita, M. Aoki, and Y. Takai. Metformin attenuates hypoxia-induced resistance to cisplatin in the hepg2 cell line. *Oncol Lett*, 17(2):2431–2440, 2019. ISSN 1792-1074 (Print) 1792-1074. 10.3892/ol.2018.9869. 137

J. Ye, K. Chen, L. Qi, R. Li, H. Tang, C. Zhou, and W. Zhai. Metformin suppresses hypoxiainduced migration via the hif1alpha/vegf pathway in gallbladder cancer in vitro and in vivo. *Oncol Rep*, 40(6):3501–3510, 2018. ISSN 1021-335x. 10.3892/or.2018.6751. 137

K. Suzuki, O. Takeuchi, Y. Suzuki, and Y. Kitagawa. Mechanisms of metformin's antitumor activity against gemcitabineresistant pancreatic adenocarcinoma. *Int J Oncol*, 54(2):764–772, 2019. ISSN 1019-6439. 10.3892/ijo.2018.4662. 137

S. H. Kim, J. W. Jeong, J. A. Park, J. W. Lee, J. H. Seo, B. K. Jung, M. K. Bae, and K. W. Kim. Regulation of the hif-1alpha stability by histone deacetylases. *Oncol Rep*, 17(3):647–51, 2007. ISSN 1021-335X (Print) 1021-335x. 138

D. M. Hutt, D. M. Roth, H. Vignaud, C. Cullin, and M. Bouchecareilh. The histone deacetylase inhibitor, vorinostat, represses hypoxia inducible factor 1 alpha expression through translational inhibition. *PLoS One*, 9(8):e106224, 2014. ISSN 1932-6203. 10.1371/journal.pone.0106224. 138